



RECEIVED

AUG 28 2000

TECH CENTER 1600/2900

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: BIRKELUND, Svend  
CHRISTIANSEN, Gunna  
HEBSGAARD PEDERSEN, Anna-Sofie  
MYGIND, Per  
KNUDSEN, Katrine

(ii) TITLE OF INVENTION: SURFACE EXPOSED PROTEINS FROM CHLAMYDIA  
PNEUMONIAE

(iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.  
(B) STREET: 624 Ninth Street, N.W., Suite 300  
(C) CITY: Washington  
(D) STATE: D.C.  
(E) COUNTRY: USA  
(F) ZIP: 20001

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 09/446,677  
(B) FILING DATE: 24-MAR-2000

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/DK98/00266  
(B) FILING DATE: 19-JUN-1998

vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: DK 0744/97  
(B) FILING DATE: 23-JUN-1997

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: COOPER, Iver P.  
(B) REGISTRATION NUMBER: 28,005  
(C) REFERENCE/DOCKET NUMBER: BIRKELUND=1

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 202-628-5197  
(B) TELEFAX: 202-737-3528

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3200 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 205...2987

(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAATGTCGAA GAGAGCACTA ACCAGGAAAA TTGCGATTTC ATAAACCCAC TTTATTATTA	60
AATTCTTACT TGCATCATAT AAAATAGAAA ACTCAGAGAG TCAAGATAAA AATTCTTGAC	120
AGCTGTTTTG TCATCTTTAA CTTGATTTAC TTATTTTGTT TCTATATTGA TGCGAATAGT	180
TCTCTAAAAA ACAAAGCAT TACC ATG AAG ACT TCG ATT CCT TGG GTT TTA	231
Met Lys Thr Ser Ile Pro Trp Val Leu	
1 5	
GTT TCC TCC GTG TTA GCT TTC TCA TGT CAC CTA CAG TCA CTA GCT AAC	279
Val Ser Ser Val Leu Ala Phe Ser Cys His Leu Gln Ser Leu Ala Asn	
10 15 20 25	
GAG GAA CTT TTA TCA CCT GAT GAT AGC TTT AAT GGA AAT ATC GAT TCA	327
Glu Glu Leu Leu Ser Pro Asp Asp Ser Phe Asn Gly Asn Ile Asp Ser	
30 35 40	
GGA ACG TTT ACT CCA AAA ACT TCA GCC ACA ACA TAT TCT CTA ACA GGA	375
Gly Thr Phe Thr Pro Lys Thr Ser Ala Thr Thr Tyr Ser Leu Thr Gly	
45 50 55	
GAT GTC TTC TTT TAC GAG CCT GGA AAA GGC ACT CCC TTA TCT GAC AGT	423
Asp Val Phe Phe Tyr Glu Pro Gly Lys Gly Thr Pro Leu Ser Asp Ser	
60 65 70	
TGT TTT AAG CAA ACC ACG GAC AAT CTT ACC TTC TTG GGG AAC GGT CAT	471
Cys Phe Lys Gln Thr Thr Asp Asn Leu Thr Phe Leu Gly Asn Gly His	
75 80 85	
AGC TTA ACG TTT GGC TTT ATA GAT GCT GGC ACT CAT GCA GGT GCT GCT	519
Ser Leu Thr Phe Gly Phe Ile Asp Ala Gly Thr His Ala Gly Ala Ala	
90 95 100 105	
GCA TCT ACA ACA GCA AAT AAG AAT CTT ACC TTC TCA GGG TTT TCC TTA	567
Ala Ser Thr Thr Ala Asn Lys Asn Leu Thr Phe Ser Gly Phe Ser Leu	
110 115 120	
CTG AGT TTT GAT TCC TCT CCT AGC ACA ACG GTT ACT ACA GGT CAG GGA	615
Leu Ser Phe Asp Ser Ser Pro Ser Thr Thr Val Thr Thr Gly Gln Gly	
125 130 135	
ACG CTT TCC TCA GCA GGA GGC GTA AAT TTA GAA AAT ATT CGT AAA CTT	663
Thr Leu Ser Ser Ala Gly Gly Val Asn Leu Glu Asn Ile Arg Lys Leu	
140 145 150	
GTA GTT GCT GGG AAT TTT TCT ACT GCA GAT GGT GGA GCT ATC AAA GGA	711
Val Val Ala Gly Asn Phe Ser Thr Ala Asp Gly Gly Ala Ile Lys Gly	
155 160 165	
GCG TCT TTC CTT TTA ACT GGC ACT TCT GGA GAT GCT CTT TTT AGT AAC	759
Ala Ser Phe Leu Leu Thr Gly Thr Ser Gly Asp Ala Leu Phe Ser Asn	

170					175					180					185	
AAC	TCT	TCA	TCA	ACA	AAG	GGA	GGA	GCA	ATT	GCT	ACT	ACA	GCA	GGC	GCT	807
Asn	Ser	Ser	Ser	Thr	Lys	Gly	Gly	Ala	Ile	Ala	Thr	Thr	Ala	Gly	Ala	
				190					195					200		
CGC	ATA	GCA	AAT	AAC	ACA	GGT	TAT	GTT	AGA	TTC	CTA	TCT	AAC	ATA	GCG	855
Arg	Ile	Ala	Asn	Asn	Thr	Gly	Tyr	Val	Arg	Phe	Leu	Ser	Asn	Ile	Ala	
			205					210					215			
TCT	ACG	TCA	GGA	GGC	GCT	ATC	GAT	GAT	GAA	GGC	ACG	TCG	ATA	CTA	TCG	903
Ser	Thr	Ser	Gly	Gly	Ala	Ile	Asp	Asp	Glu	Gly	Thr	Ser	Ile	Leu	Ser	
		220					225					230				
AAC	AAC	AAA	TTT	CTA	TAT	TTT	GAA	GGG	AAT	GCA	GCG	AAA	ACT	ACT	GGC	951
Asn	Asn	Lys	Phe	Leu	Tyr	Phe	Glu	Gly	Asn	Ala	Ala	Lys	Thr	Thr	Gly	
	235					240				245						
GGT	GCG	ATC	TGC	AAC	ACC	AAG	GCG	AGT	GGA	TCT	CCT	GAA	CTG	ATA	ATC	999
Gly	Ala	Ile	Cys	Asn	Thr	Lys	Ala	Ser	Gly	Ser	Pro	Glu	Leu	Ile	Ile	
250					255				260						265	
TCT	AAC	AAT	AAG	ACT	CTG	ATC	TTT	GCT	TCA	AAC	GTA	GCA	GAA	ACA	AGC	1047
Ser	Asn	Asn	Lys	Thr	Leu	Ile	Phe	Ala	Ser	Asn	Val	Ala	Glu	Thr	Ser	
				270				275						280		
GGT	GCG	GCC	ATC	CAT	GCT	AAA	AAG	CTA	GCC	CTT	TCC	TCT	GGA	GGC	TTT	1095
Gly	Gly	Ala	Ile	His	Ala	Lys	Lys	Leu	Ala	Leu	Ser	Ser	Gly	Gly	Phe	
			285					290					295			
ACA	GAG	TTT	CTA	CGA	AAT	AAT	GTC	TCA	TCA	GCA	ACT	CCT	AAG	GGG	GGT	1143
Thr	Glu	Phe	Leu	Arg	Asn	Asn	Val	Ser	Ser	Ala	Thr	Pro	Lys	Gly	Gly	
		300					305					310				
GCT	ATC	AGC	ATC	GAT	GCC	TCA	GGA	GAG	CTC	AGT	CTT	TCT	GCA	GAG	ACA	1191
Ala	Ile	Ser	Ile	Asp	Ala	Ser	Gly	Glu	Leu	Ser	Leu	Ser	Ala	Glu	Thr	
	315					320					325					
GGA	AAC	ATT	ACC	TTT	GTA	AGA	AAT	ACC	CTT	ACA	ACA	ACC	GGA	AGT	ACC	1239
Gly	Asn	Ile	Thr	Phe	Val	Arg	Asn	Thr	Leu	Thr	Thr	Thr	Gly	Ser	Thr	
330					335				340						345	
GAT	ACT	CCT	AAA	CGT	AAT	GCG	ATC	AAC	ATA	GGA	AGT	AAC	GGG	AAA	TTC	1287
Asp	Thr	Pro	Lys	Arg	Asn	Ala	Ile	Asn	Ile	Gly	Ser	Asn	Gly	Lys	Phe	
				350				355						360		
ACG	GAA	TTA	CGG	GCT	GCT	AAA	AAT	CAT	ACA	ATT	TTC	TTC	TAT	GAT	CCC	1335
Thr	Glu	Leu	Arg	Ala	Ala	Lys	Asn	His	Thr	Ile	Phe	Phe	Tyr	Asp	Pro	
			365				370						375			
ATC	ACT	TCA	GAA	GGA	ACC	TCA	TCA	GAC	GTA	TTG	AAG	ATA	AAT	AAC	GGC	1383
Ile	Thr	Ser	Glu	Gly	Thr	Ser	Ser	Asp	Val	Leu	Lys	Ile	Asn	Asn	Gly	
		380				385						390				
TCT	GCG	GGA	GCT	CTC	AAT	CCA	TAT	CAA	GGA	ACG	ATT	CTA	TTT	TCT	GGA	1431
Ser	Ala	Gly	Ala	Leu	Asn	Pro	Tyr	Gln	Gly	Thr	Ile	Leu	Phe	Ser	Gly	
	395					400					405					

GAA ACC CTA ACA GCA GAT GAA CTT AAA GTT GCT GAC AAT TTA AAA TCT Glu Thr Leu Thr Ala Asp Glu Leu Lys Val Ala Asp Asn Leu Lys Ser 410 415 420 425	1479
TCA TTC ACG CAG CCA GTC TCC CTA TCC GGA GGA AAG TTA TTG CTA CAA Ser Phe Thr Gln Pro Val Ser Leu Ser Gly Gly Lys Leu Leu Leu Gln 430 435 440	1527
AAG GGA GTC ACT TTA GAG AGC ACG AGC TTC TCT CAA GAG GCC GGT TCT Lys Gly Val Thr Leu Glu Ser Thr Ser Phe Ser Gln Glu Ala Gly Ser 445 450 455	1575
CTC CTC GGC ATG GAT TCA GGA ACG ACA TTA TCA ACT ACA GCT GGG AGT Leu Leu Gly Met Asp Ser Gly Thr Thr Leu Ser Thr Thr Ala Gly Ser 460 465 470	1623
ATT ACA ATC ACG AAC CTA GGA ATC AAT GTT GAC TCC TTA GGT CTT AAG Ile Thr Ile Thr Asn Leu Gly Ile Asn Val Asp Ser Leu Gly Leu Lys 475 480 485	1671
CAG CCC GTC AGC CTA ACA GCA AAA GGT GCT TCA AAT AAA GTG ATC GTA Gln Pro Val Ser Leu Thr Ala Lys Gly Ala Ser Asn Lys Val Ile Val 490 495 500 505	1719
TCT GGG AAG CTC AAC CTG ATT GAT ATT GAA GGG AAC ATT TAT GAA AGT Ser Gly Lys Leu Asn Leu Ile Asp Ile Glu Gly Asn Ile Tyr Glu Ser 510 515 520	1767
CAT ATG TTC AGC CAT GAC CAG CTC TTC TCT CTA TTA AAA ATC ACG GTT His Met Phe Ser His Asp Gln Leu Phe Ser Leu Leu Lys Ile Thr Val 525 530 535	1815
GAT GCT GAT GTT GAT ACT AAC GTT GAC ATC AGC AGC CTT ATC CCT GTT Asp Ala Asp Val Asp Thr Asn Val Asp Ile Ser Ser Leu Ile Pro Val 540 545 550	1863
CCT GCT GAG GAT CCT AAT TCA GAA TAC GGA TTC CAA GGA CAA TGG AAT Pro Ala Glu Asp Pro Asn Ser Glu Tyr Gly Phe Gln Gly Gln Trp Asn 555 560 565	1911
GTT AAT TGG ACT ACG GAT ACA GCT ACA AAT ACA AAA GAG GCC ACG GCA Val Asn Trp Thr Thr Asp Thr Ala Thr Asn Thr Lys Glu Ala Thr Ala 570 575 580 585	1959
ACT TGG ACC AAA ACA GGA TTT GTT CCC AGC CCC GAA AGA AAA TCT GCG Thr Trp Thr Lys Thr Gly Phe Val Pro Ser Pro Glu Arg Lys Ser Ala 590 595 600	2007
TTA GTA TGC AAT ACC CTA TGG GGA GTC TTT ACT GAC ATT CGC TCT CTG Leu Val Cys Asn Thr Leu Trp Gly Val Phe Thr Asp Ile Arg Ser Leu 605 610 615	2055
CAA CAG CTT GTA GAG ATC GGC GCA ACT GGT ATG GAA CAC AAA CAA GGT Gln Gln Leu Val Glu Ile Gly Ala Thr Gly Met Glu His Lys Gln Gly 620 625 630	2103
TTC TGG GTT TCC TCC ATG ACG AAC TTC CTG CAT AAG ACT GGA GAT GAA Phe Trp Val Ser Ser Met Thr Asn Phe Leu His Lys Thr Gly Asp Glu	2151

635	640	645	
AAT CGC AAA GGC TTC CGT CAT ACC TCT GGA GGC TAC GTC ATC GGT GGA Asn Arg Lys Gly Phe Arg His Thr Ser Gly Gly Tyr Val Ile Gly Gly 650 655 660 665	2199		
AGT GCT CAC ACT CCT AAA GAC GAC CTA TTT ACC TTT GCG TTC TGC CAT Ser Ala His Thr Pro Lys Asp Asp Leu Phe Thr Phe Ala Phe Cys His 670 675 680	2247		
CTC TTT GCT AGA GAC AAA GAT TGT TTT ATC GCT CAC AAC AAC TCT AGA Leu Phe Ala Arg Asp Lys Asp Cys Phe Ile Ala His Asn Asn Ser Arg 685 690 695	2295		
ACC TAC GGT GGA ACT TTA TTC TTC AAG CAC TCT CAT ACC CTA CAA CCC Thr Tyr Gly Gly Thr Leu Phe Phe Lys His Ser His Thr Leu Gln Pro 700 705 710	2343		
CAA AAC TAT TTG AGA TTA GGA AGA GCA AAG TTT TCT GAA TCA GCT ATA Gln Asn Tyr Leu Arg Leu Gly Arg Ala Lys Phe Ser Glu Ser Ala Ile 715 720 725	2391		
GAA AAA TTC CCT AGG GAA ATT CCC CTA GCC TTG GAT GTC CAA GTT TCG Glu Lys Phe Pro Arg Glu Ile Pro Leu Ala Leu Asp Val Gln Val Ser 730 735 740 745	2439		
TTC AGC CAT TCA GAC AAC CGT ATG GAA ACG CAC TAT ACC TCA TTG CCA Phe Ser His Ser Asp Asn Arg Met Glu Thr His Tyr Thr Ser Leu Pro 750 755 760	2487		
GAA TCC GAA GGT TCT TGG AGC AAC GAG TGT ATA GCT GGT GGT ATC GGC Glu Ser Glu Gly Ser Trp Ser Asn Glu Cys Ile Ala Gly Gly Ile Gly 765 770 775	2535		
CTA GAC CTT CCT TTT GTT CTT TCC AAC CCA CAT CCT CTT TTC AAG ACC Leu Asp Leu Pro Phe Val Leu Ser Asn Pro His Pro Leu Phe Lys Thr 780 785 790	2583		
TTC ATT CCA CAG ATG AAA GTC GAA ATG GTT TAT GTA TCA CAA AAT AGC Phe Ile Pro Gln Met Lys Val Glu Met Val Tyr Val Ser Gln Asn Ser 795 800 805	2631		
TTC TTC GAA AGC TCT AGT GAT GGC CGT GGT TTT AGT ATT GGA AGG CTG Phe Phe Glu Ser Ser Ser Asp Gly Arg Gly Phe Ser Ile Gly Arg Leu 810 815 820 825	2679		
CTT AAC CTC TCG ATT CCT GTG GGT GCG AAA TTC GTG CAG GGG GAT ATC Leu Asn Leu Ser Ile Pro Val Gly Ala Lys Phe Val Gln Gly Asp Ile 830 835 840	2727		
GGA GAT TCC TAC ACC TAT GAT CTC TCA GGA TTC TTT GTT TCC GAT GTC Gly Asp Ser Tyr Thr Tyr Asp Leu Ser Gly Phe Phe Val Ser Asp Val 845 850 855	2775		
TAT CGT AAC AAT CCC CAA TCT ACA GCG ACT CTT GTG ATG AGC CCA GAC Tyr Arg Asn Asn Pro Gln Ser Thr Ala Thr Leu Val Met Ser Pro Asp 860 865 870	2823		

```

TCT TGG AAA ATT CGC GGT GGC AAT CTT TCA AGA CAG GCA TTT TTA CTG      2871
Ser Trp Lys Ile Arg Gly Gly Asn Leu Ser Arg Gln Ala Phe Leu Leu
      875                      880                      885

AGG GGT AGC AAC AAC TAC GTC TAC AAC TCC AAT TGT GAG CTC TTC GGA      2919
Arg Gly Ser Asn Asn Tyr Val Tyr Asn Ser Asn Cys Glu Leu Phe Gly
      890                      895                      900                      905

CAT TAC GCT ATG GAA CTC CGT GGA TCT TCA AGG AAC TAC AAT GTA GAT      2967
His Tyr Ala Met Glu Leu Arg Gly Ser Ser Arg Asn Tyr Asn Val Asp
              910                      915                      920

GTT GGT ACC AAA CTC CGA TT CTAGATTGCT AAAACTCCCT AGTTCTTCTA GGGAG      3022
Val Gly Thr Lys Leu Arg Phe
              925

TTTTCTCATA CTTTATAGGGA AATATTTGCT ATAGGGAATG CTTTCCTTGC AAACGTGTAAG      3082
AAATAACATT TGTCCCTCTT CAAAAAAGAT TTCTTTTAAT AATTTCTAGT TATAATTTTA      3142
TTTTAAAAAC AGTTAAATAA TTAATAGACA ATAATCTATT CTTATTGACT TCTTTTTTT      3200

```

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 928 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Lys Thr Ser Ile Pro Trp Val Leu Val Ser Ser Val Leu Ala Phe
  1                      5                      10                      15

Ser Cys His Leu Gln Ser Leu Ala Asn Glu Glu Leu Leu Ser Pro Asp
      20                      25                      30

Asp Ser Phe Asn Gly Asn Ile Asp Ser Gly Thr Phe Thr Pro Lys Thr
      35                      40                      45

Ser Ala Thr Thr Tyr Ser Leu Thr Gly Asp Val Phe Phe Tyr Glu Pro
      50                      55                      60

Gly Lys Gly Thr Pro Leu Ser Asp Ser Cys Phe Lys Gln Thr Thr Asp
      65                      70                      75                      80

Asn Leu Thr Phe Leu Gly Asn Gly His Ser Leu Thr Phe Gly Phe Ile
      85                      90                      95

Asp Ala Gly Thr His Ala Gly Ala Ala Ala Ser Thr Thr Ala Asn Lys
      100                      105                      110
Asn Leu Thr Phe Ser Gly Phe Ser Leu Leu Ser Phe Asp Ser Ser Pro
      115                      120                      125

Ser Thr Thr Val Thr Thr Gly Gln Gly Thr Leu Ser Ser Ala Gly Gly

```

130	135	140
Val Asn Leu Glu Asn Ile Arg Lys Leu Val Val Ala Gly Asn Phe Ser		
145	150	155 160
Thr Ala Asp Gly Gly Ala Ile Lys Gly Ala Ser Phe Leu Leu Thr Gly		
	165	170 175
Thr Ser Gly Asp Ala Leu Phe Ser Asn Asn Ser Ser Ser Thr Lys Gly		
	180	185 190
Gly Ala Ile Ala Thr Thr Ala Gly Ala Arg Ile Ala Asn Asn Thr Gly		
	195	200 205
Tyr Val Arg Phe Leu Ser Asn Ile Ala Ser Thr Ser Gly Gly Ala Ile		
	210	215 220
Asp Asp Glu Gly Thr Ser Ile Leu Ser Asn Asn Lys Phe Leu Tyr Phe		
	225	230 235 240
Glu Gly Asn Ala Ala Lys Thr Thr Gly Gly Ala Ile Cys Asn Thr Lys		
	245	250 255
Ala Ser Gly Ser Pro Glu Leu Ile Ile Ser Asn Asn Lys Thr Leu Ile		
	260	265 270
Phe Ala Ser Asn Val Ala Glu Thr Ser Gly Gly Ala Ile His Ala Lys		
	275	280 285
Lys Leu Ala Leu Ser Ser Gly Gly Phe Thr Glu Phe Leu Arg Asn Asn		
	290	295 300
Val Ser Ser Ala Thr Pro Lys Gly Gly Ala Ile Ser Ile Asp Ala Ser		
	305	310 315 320
Gly Glu Leu Ser Leu Ser Ala Glu Thr Gly Asn Ile Thr Phe Val Arg		
	325	330 335
Asn Thr Leu Thr Thr Thr Gly Ser Thr Asp Thr Pro Lys Arg Asn Ala		
	340	345 350
Ile Asn Ile Gly Ser Asn Gly Lys Phe Thr Glu Leu Arg Ala Ala Lys		
	355	360 365
Asn His Thr Ile Phe Phe Tyr Asp Pro Ile Thr Ser Glu Gly Thr Ser		
	370	375 380
Ser Asp Val Leu Lys Ile Asn Asn Gly Ser Ala Gly Ala Leu Asn Pro		
	385	390 395 400
Tyr Gln Gly Thr Ile Leu Phe Ser Gly Glu Thr Leu Thr Ala Asp Glu		
	405	410 415
Leu Lys Val Ala Asp Asn Leu Lys Ser Ser Phe Thr Gln Pro Val Ser		
	420	425 430
Leu Ser Gly Gly Lys Leu Leu Leu Gln Lys Gly Val Thr Leu Glu Ser		
	435	440 445

Thr	Ser	Phe	Ser	Gln	Glu	Ala	Gly	Ser	Leu	Leu	Gly	Met	Asp	Ser	Gly	
450						455					460					
Thr	Thr	Leu	Ser	Thr	Thr	Ala	Gly	Ser	Ile	Thr	Ile	Thr	Asn	Leu	Gly	
465					470					475					480	
Ile	Asn	Val	Asp	Ser	Leu	Gly	Leu	Lys	Gln	Pro	Val	Ser	Leu	Thr	Ala	
				485					490						495	
Lys	Gly	Ala	Ser	Asn	Lys	Val	Ile	Val	Ser	Gly	Lys	Leu	Asn	Leu	Ile	
			500					505					510			
Asp	Ile	Glu	Gly	Asn	Ile	Tyr	Glu	Ser	His	Met	Phe	Ser	His	Asp	Gln	
		515					520					525				
Leu	Phe	Ser	Leu	Leu	Lys	Ile	Thr	Val	Asp	Ala	Asp	Val	Asp	Thr	Asn	
	530					535					540					
Val	Asp	Ile	Ser	Ser	Leu	Ile	Pro	Val	Pro	Ala	Glu	Asp	Pro	Asn	Ser	
545					550					555					560	
Glu	Tyr	Gly	Phe	Gln	Gly	Gln	Trp	Asn	Val	Asn	Trp	Thr	Thr	Asp	Thr	
				565					570					575		
Ala	Thr	Asn	Thr	Lys	Glu	Ala	Thr	Ala	Thr	Trp	Thr	Lys	Thr	Gly	Phe	
			580					585						590		
Val	Pro	Ser	Pro	Glu	Arg	Lys	Ser	Ala	Leu	Val	Cys	Asn	Thr	Leu	Trp	
		595					600					605				
Gly	Val	Phe	Thr	Asp	Ile	Arg	Ser	Leu	Gln	Gln	Leu	Val	Glu	Ile	Gly	
	610					615					620					
Ala	Thr	Gly	Met	Glu	His	Lys	Gln	Gly	Phe	Trp	Val	Ser	Ser	Met	Thr	
625					630					635					640	
Asn	Phe	Leu	His	Lys	Thr	Gly	Asp	Glu	Asn	Arg	Lys	Gly	Phe	Arg	His	
				645					650					655		
Thr	Ser	Gly	Gly	Tyr	Val	Ile	Gly	Gly	Ser	Ala	His	Thr	Pro	Lys	Asp	
			660					665					670			
Asp	Leu	Phe	Thr	Phe	Ala	Phe	Cys	His	Leu	Phe	Ala	Arg	Asp	Lys	Asp	
		675					680					685				
Cys	Phe	Ile	Ala	His	Asn	Asn	Ser	Arg	Thr	Tyr	Gly	Gly	Thr	Leu	Phe	
	690					695					700					
Phe	Lys	His	Ser	His	Thr	Leu	Gln	Pro	Gln	Asn	Tyr	Leu	Arg	Leu	Gly	
705					710					715					720	
Arg	Ala	Lys	Phe	Ser	Glu	Ser	Ala	Ile	Glu	Lys	Phe	Pro	Arg	Glu	Ile	
				725					730					735		
Pro	Leu	Ala	Leu	Asp	Val	Gln	Val	Ser	Phe	Ser	His	Ser	Asp	Asn	Arg	
			740					745					750			



Met Glu Thr His Tyr Thr Ser Leu Pro Glu Ser Glu Gly Ser Trp Ser  
755 760 765

Asn Glu Cys Ile Ala Gly Gly Ile Gly Leu Asp Leu Pro Phe Val Leu  
770 775 780

Ser Asn Pro His Pro Leu Phe Lys Thr Phe Ile Pro Gln Met Lys Val  
785 790 795 800

Glu Met Val Tyr Val Ser Gln Asn Ser Phe Phe Glu Ser Ser Ser Asp  
805 810 815

Gly Arg Gly Phe Ser Ile Gly Arg Leu Leu Asn Leu Ser Ile Pro Val  
820 825 830

Gly Ala Lys Phe Val Gln Gly Asp Ile Gly Asp Ser Tyr Thr Tyr Asp  
835 840 845

Leu Ser Gly Phe Phe Val Ser Asp Val Tyr Arg Asn Asn Pro Gln Ser  
850 855 860

Thr Ala Thr Leu Val Met Ser Pro Asp Ser Trp Lys Ile Arg Gly Gly  
865 870 875 880

Asn Leu Ser Arg Gln Ala Phe Leu Leu Arg Gly Ser Asn Asn Tyr Val  
885 890 895

Tyr Asn Ser Asn Cys Glu Leu Phe Gly His Tyr Ala Met Glu Leu Arg  
900 905 910

Gly Ser Ser Arg Asn Tyr Asn Val Asp Val Gly Thr Lys Leu Arg Phe  
915 920 925

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2815 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAAATCGC	AATTTTCCTG	GTTAGTGCTC	TCTTCGACAT	TGGCATGTTT	TACTAGTTGT	60
TCCACTGTTT	TTGCTGCAAC	TGCTGAAAAT	ATAGGCCCCCT	CTGATAGCTT	TGACGGAAGT	120
ACTAACACAG	GCACCTATAC	TCCTAAAAAT	ACGACTACTG	GAATAGACTA	TACTCTGACA	180
GGAGATATAA	CTCTGCAAAA	CCTTGGGGAT	TCGGCAGCTT	TAACGAAGGG	TTGTTTTTCT	240
GACACTACGG	AATCTTTAAG	CTTTGCCGGT	AAGGGGTACT	CACTTTCTTT	TTTAAATATT	300
AAGTCTAGTG	CTGAAGGCGC	AGCACTTTCT	GTTACAACCTG	ATAAAAATCT	GTCGCTAACA	360
GGATTTTCGA	GTCTTACTTT	CTTAGCGGCC	CCATCATCGG	TAATCACAAC	CCCCTCAGGA	420
AAAGGTGCAG	TTAAATGTGG	AGGGGATCTT	ACATTTGATA	ACAATGGAAC	TATTTTATTT	480
AAACAAGATT	ACTGTGAGGA	AAATGGCGGA	GCCATTTCTA	CCAAGAATCT	TTCTTTGAAA	540
AACAGCACGG	GATCGATTTT	TTTTGAAGGG	AATAAATCGA	GCGCAACAGG	GAAAAAAGGT	600
GGGGCTATTT	GTGCTACTGG	TACTGTAGAT	ATTACAAATA	ATACGGCTCC	TACCCTCTTC	660
TCGAACAATA	TTGCTGAAGC	TGCAGGTGGA	GCTATAAATA	GCACAGGAAA	CTGTACAATT	720
ACAGGGAATA	CGTCTCTTGT	ATTTTCTGAA	AATAGTGTGA	CAGCGACCGC	AGGAAATGGA	780

```

GGAGCTCTTT CTGGAGATGC CGATGTTACC ATATCTGGGA ATCAGAGTGT AACTTTCTCA      840
GGAAACCAAG CTGTAGCTAA TGGCGGAGCC ATTTATGCTA AGAAGCTTAC ACTGGCTTCC      900
GGGGGGGGGG GGGGTATCTC CTTTTCTAAC AATATAGTCC AAGGTACCAC TGCAGGTAAT      960
GGTGGAGCCA TTTCTATACT GGCAGCTGGA GAGTGTAGTC TTTCAGCAGA AGCAGGGGAC     1020
ATTACCTTCA ATGGGAATGC CATTGTTGCA ACTACACCAC AAAC TACAAA AAGAAATTCT     1080
ATTGACATAG GATCTACTGC AAAGATCACG AATTTACGTG CAATATCTGG GCATAGCATC     1140
TTTTTCTACG ATCCGATTAC TGCTAATACG GCTGCGGATT CTACAGATAC TTAAATCTC      1200
AATAAGGCTG ATGCAGGTAA TAGTACAGAT TATAGTGGGT CGATTGTTTT TTCTGGTGAA      1260
AAGCTCTCTG AAGATGAAGC AAAAGTTGCA GACAACCTCA CTTCTACGCT GAAGCAGCCT     1320
GTAAC TCTAA CTGCAGGAAA TTTAGTACTT AAACGTGGTG TCACTCTCGA TACGAAAGGC     1380
TTTACTCAGA CCGCGGGTTC CTCTGTTATT ATGGATGCGG GCACAACGTT AAAAGCAAGT     1440
ACAGAGGAGG TCACTTTAAC AGGTCTTTCC ATTCCTGTAG ACTCTTTAGG CGAGGGTAAG     1500
AAAGTTGTAA TTGCTGCTTC TGCAGCAAGT AAAAATGTAG CCCTTAGTGG TCCGATTCTT     1560
CTTTTGGATA ACCAAGGGAA TGCTTATGAA AATCACGACT TAGGAAAAAC TCAAGACTTT     1620
TCATTTGTGC AGCTCTCTGC TCTGGGTA CTGCAACACTA CAGATGTTCC AGCGGTTTCT     1680
ACAGTAGCAA CTCCTACGCA CTATGGGTAT CAAGGTACTT GGGGAATGAC TTGGGTTGAT     1740
GATACCGCAA GCACTCCAAA GACTAAGACA GCGACATTAG CTTGGACCAA TACAGGCTAC     1800
CTTCCGAATC CTGAGCGTCA AGGACCTTTA GTTCCTAATA GCCTTTGGGG ATCTTTTTCA     1860
GACATCCAAG CGATTCAAGG TGTCATAGAG AGAAGTGCTT TGACTCTTTG TTCAGATCGA     1920
GGCTTCTGGG CTGCGGGAGT CGCCAATTTC TTAGATAAAG ATAAGAAAGG GGAAAAACGC     1980
AAATACCGTC ATAAATCTGG TGGATATGCT ATCGGAGGTG CAGCGCAAAC TTGTTCTGAA     2040
AACTTAATTA GCTTTGCCTT TTGCCAACTC TTTGGTAGCG ATAAAGATTT CTTAGTCGCT     2100
AAAAATCATA CTGATACCTA TGCAGGAGCC TTCTATATCC AACACATTAC AGAATGTAGT     2160
GGGTTTCATAG GTTGTCTCTT AGATAAACTT CCTGGCTCTT GGAGTCATAA ACCCCTCGTT     2220
TTAGAAGGGG AGCTCGCTTA TAGCCACGTC AGTAATGATC TGAAGACAAA GTATACTGCG     2280
TATCCTGAGG TGAAGGTTT TGGGGGAAT AATGCTTTTA ACATGATGTT GGGAGCTTCT     2340
TCTCATTTCT ATCCTGAATA CCTGCATTGT TTTGATACCT ATGCTCCATA CATCAAAC TG     2400
AATCTGACCT ATATACGTCA GGACAGCTTC TCGGAGAAAG GTACAGAAGG AAGATCTTTT     2460
GATGACAGCA ACCTCTTCAA TTTATCTTTG CCTATAGGGG TGAAGTTTGA GAAGTTCTCT     2520
GATTGTAATG ACTTTTCTTA TGATCTGACT TTATCCTATG TTCCTGATCT TATCCGCAAT     2580
GATCCCAAAT GCACTACAGC ACTTGTAATC AGCGGAGCCT CTTGGGAAAC TTATGCCAAT     2640
AACTTAGCAC GACAGGCCTT GCAAGTGCGT GCAGGCAGTC ACTACGCCTT CTCTCCTATG     2700
TTTGAAGTGC TCGGCCAGTT TGTCTTTGAA GTTCGTGGAT CCTCACGGAT TTATAATGTA     2760
GATCTTGGGG GTAAGTTCCA ATTCTAGGAG CGTCTCTCAT GTCTCAGAAA TTCTG      2815

```

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 928 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Lys Ser Gln Phe Ser Trp Leu Val Leu Ser Ser Thr Leu Ala Cys
 1              5              10              15

Phe Thr Ser Cys Ser Thr Val Phe Ala Ala Thr Ala Glu Asn Ile Gly
      20              25              30

Pro Ser Asp Ser Phe Asp Gly Ser Thr Asn Thr Gly Thr Tyr Thr Pro
      35              40              45

Lys Asn Thr Thr Thr Gly Ile Asp Tyr Thr Leu Thr Gly Asp Ile Thr
      50              55              60

```

Leu	Gln	Asn	Leu	Gly	Asp	Ser	Ala	Ala	Leu	Thr	Lys	Gly	Cys	Phe	Ser	
65					70					75					80	
Asp	Thr	Thr	Glu	Ser	Leu	Ser	Phe	Ala	Gly	Lys	Gly	Tyr	Ser	Leu	Ser	
				85					90					95		
Phe	Leu	Asn	Ile	Lys	Ser	Ser	Ala	Glu	Gly	Ala	Ala	Leu	Ser	Val	Thr	
			100					105					110			
Thr	Asp	Lys	Asn	Leu	Ser	Leu	Thr	Gly	Phe	Ser	Ser	Leu	Thr	Phe	Leu	
		115					120					125				
Ala	Ala	Pro	Ser	Ser	Val	Ile	Thr	Thr	Pro	Ser	Gly	Lys	Gly	Ala	Val	
	130					135					140					
Lys	Cys	Gly	Gly	Asp	Leu	Thr	Phe	Asp	Asn	Asn	Gly	Thr	Ile	Leu	Phe	
145					150				155					160		
Lys	Gln	Asp	Tyr	Cys	Glu	Glu	Asn	Gly	Gly	Ala	Ile	Ser	Thr	Lys	Asn	
				165				170						175		
Leu	Ser	Leu	Lys	Asn	Ser	Thr	Gly	Ser	Ile	Ser	Phe	Glu	Gly	Asn	Lys	
			180					185					190			
Ser	Ser	Ala	Thr	Gly	Lys	Lys	Gly	Gly	Ala	Ile	Cys	Ala	Thr	Gly	Thr	
		195					200					205				
Val	Asp	Ile	Thr	Asn	Asn	Thr	Ala	Pro	Thr	Leu	Phe	Ser	Asn	Asn	Ile	
	210					215					220					
Ala	Glu	Ala	Ala	Gly	Gly	Ala	Ile	Asn	Ser	Thr	Gly	Asn	Cys	Thr	Ile	
225					230					235					240	
Thr	Gly	Asn	Thr	Ser	Leu	Val	Phe	Ser	Glu	Asn	Ser	Val	Thr	Ala	Thr	
				245					250					255		
Ala	Gly	Asn	Gly	Gly	Ala	Leu	Ser	Gly	Asp	Ala	Asp	Val	Thr	Ile	Ser	
			260					265				270				
Gly	Asn	Gln	Ser	Val	Thr	Phe	Ser	Gly	Asn	Gln	Ala	Val	Ala	Asn	Gly	
	275						280					285				
Gly	Ala	Ile	Tyr	Ala	Lys	Lys	Leu	Thr	Leu	Ala	Ser	Gly	Gly	Gly	Gly	
	290					295					300					
Gly	Ile	Ser	Phe	Ser	Asn	Asn	Ile	Val	Gln	Gly	Thr	Thr	Ala	Gly	Asn	
305					310					315				320		
Gly	Gly	Ala	Ile	Ser	Ile	Leu	Ala	Ala	Gly	Glu	Cys	Ser	Leu	Ser	Ala	
				325					330				335			
Glu	Ala	Gly	Asp	Ile	Thr	Phe	Asn	Gly	Asn	Ala	Ile	Val	Ala	Thr	Thr	
			340					345					350			
Pro	Gln	Thr	Thr	Lys	Arg	Asn	Ser	Ile	Asp	Ile	Gly	Ser	Thr	Ala	Lys	
		355					360					365				

Ile Thr Asn Leu Arg Ala Ile Ser Gly His Ser Ile Phe Phe Tyr Asp  
 370 375 380  
 Pro Ile Thr Ala Asn Thr Ala Ala Asp Ser Thr Asp Thr Leu Asn Leu  
 385 390 395 400  
 Asn Lys Ala Asp Ala Gly Asn Ser Thr Asp Tyr Ser Gly Ser Ile Val  
 405 410 415  
 Phe Ser Gly Glu Lys Leu Ser Glu Asp Glu Ala Lys Val Ala Asp Asn  
 420 425 430  
 Leu Thr Ser Thr Leu Lys Gln Pro Val Thr Leu Thr Ala Gly Asn Leu  
 435 440 445  
 Val Leu Lys Arg Gly Val Thr Leu Asp Thr Lys Gly Phe Thr Gln Thr  
 450 455 460  
 Ala Gly Ser Ser Val Ile Met Asp Ala Gly Thr Thr Leu Lys Ala Ser  
 465 470 475 480  
 Thr Glu Glu Val Thr Leu Thr Gly Leu Ser Ile Pro Val Asp Ser Leu  
 485 490 495  
 Gly Glu Gly Lys Lys Val Val Ile Ala Ala Ser Ala Ala Ser Lys Asn  
 500 505 510  
 Val Ala Leu Ser Gly Pro Ile Leu Leu Leu Asp Asn Gln Gly Asn Ala  
 515 520 525  
 Tyr Glu Asn His Asp Leu Gly Lys Thr Gln Asp Phe Ser Phe Val Gln  
 530 535 540  
 Leu Ser Ala Leu Gly Thr Ala Thr Thr Thr Asp Val Pro Ala Val Pro  
 545 550 555 560  
 Thr Val Ala Thr Pro Thr His Tyr Gly Tyr Gln Gly Thr Trp Gly Met  
 565 570 575  
 Thr Trp Val Asp Asp Thr Ala Ser Thr Pro Lys Thr Lys Thr Ala Thr  
 580 585 590  
 Leu Ala Trp Thr Asn Thr Gly Tyr Leu Pro Asn Pro Glu Arg Gln Gly  
 595 600 605  
 Pro Leu Val Pro Asn Ser Leu Trp Gly Ser Phe Ser Asp Ile Gln Ala  
 610 615 620  
 Ile Gln Gly Val Ile Glu Arg Ser Ala Leu Thr Leu Cys Ser Asp Arg  
 625 630 635 640  
 Gly Phe Trp Ala Ala Gly Val Ala Asn Phe Leu Asp Lys Asp Lys Lys  
 645 650 655  
 Gly Glu Lys Arg Lys Tyr Arg His Lys Ser Gly Gly Tyr Ala Ile Gly  
 660 665 670  
 Gly Ala Ala Gln Thr Cys Ser Glu Asn Leu Ile Ser Phe Ala Phe Cys

675		680		685
Gln Leu Phe Gly Ser Asp Lys Asp Phe Leu Val Ala Lys Asn His Thr				
690		695		700
Asp Thr Tyr Ala Gly Ala Phe Tyr Ile Gln His Ile Thr Glu Cys Ser				
705		710		715
Gly Phe Ile Gly Cys Leu Leu Asp Lys Leu Pro Gly Ser Trp Ser His				
		725		730
Lys Pro Leu Val Leu Glu Gly Gln Leu Ala Tyr Ser His Val Ser Asn				
		740		745
Asp Leu Lys Thr Lys Tyr Thr Ala Tyr Pro Glu Val Lys Gly Ser Trp				
		755		760
Gly Asn Asn Ala Phe Asn Met Met Leu Gly Ala Ser Ser His Ser Tyr				
		770		775
Pro Glu Tyr Leu His Cys Phe Asp Thr Tyr Ala Pro Tyr Ile Lys Leu				
		785		790
Asn Leu Thr Tyr Ile Arg Gln Asp Ser Phe Ser Glu Lys Gly Thr Glu				
		805		810
Gly Arg Ser Phe Asp Asp Ser Asn Leu Phe Asn Leu Ser Leu Pro Ile				
		820		825
Gly Val Lys Phe Glu Lys Phe Ser Asp Cys Asn Asp Phe Ser Tyr Asp				
		835		840
Leu Thr Leu Ser Tyr Val Pro Asp Leu Ile Arg Asn Asp Pro Lys Cys				
		850		855
Thr Thr Ala Leu Val Ile Ser Gly Ala Ser Trp Glu Thr Tyr Ala Asn				
		865		870
Asn Leu Ala Arg Gln Ala Leu Gln Val Arg Ala Gly Ser His Tyr Ala				
		885		890
Phe Ser Pro Met Phe Glu Val Leu Gly Gln Phe Val Phe Glu Val Arg				
		900		905
Gly Ser Ser Arg Ile Tyr Asn Val Asp Leu Gly Gly Lys Phe Gln Phe				
		915		920
				925

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3052 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATGCGATTTT	CGCTCTGCGG	ATTTCTCTA	GTTTTTCTT	TAACATTGCT	CTCAGTCTTC	60
GACACTTCTT	TGAGTGCTAC	TACGATTTCT	TTAACCCAG	AAGATAGTTT	TCATGGAGAT	120
AGTCAGAATG	CAGAACGTTT	TTATAATGTT	CAAGCTGGGG	ATGTCTATAG	CCTTACTGGT	180
GATGTCTCAA	TATCTAACGT	CGATAACTCT	GCATTAAATA	AAGCCTGCTT	CAATGTGACC	240
TCAGGAAGTG	TGACGTTTCG	AGGAAATCAT	CATGGGTAT	ATTTTAATAA	TATTTCTCTA	300
GGAACACAA	AGGAAGGGGG	TGTACTTTGT	TGCCAAGATC	CTCAAGCAAC	GGCACGTTTT	360
TCTGGGTTCT	CCACGCTCTC	TTTTATTCAG	AGCCCGGAG	ATATTAAAGA	ACAGGGATGT	420
CTCTATTCAA	AAAATGCACT	TATGCTCTTA	AACAATTATG	TAGTGCGTTT	TGAACAAAAC	480
CAAAGTAAGA	CTAAAGGCGG	AGCTATTAGT	GGGGCGAATG	TTACTATAGT	AGGCAACTAC	540
GATTCGGTCT	CTTTCTATCA	GAATGCAGCC	ACTTTTGGAG	GTGCTATCCA	TTCTTCAGGT	600
CCCCTACAGA	TTGCAGTAAA	TCAGGCAGAG	ATAAGATTTG	CACAAAATAC	TGCCAAGAAT	660
GGTTCTGGAG	GGGCTTTGTA	CTCCGATGGT	GATATTGATA	TTGATCAGAA	TGCTTATGTT	720
CTATTTTCGAG	AAAATGAGGC	ATTGACTACT	GCTATAGGTA	AGGGAGGGGC	TGTCTGTTGT	780
CTTCCCCTT	CAGGAAGTAG	TACTCCAGTT	CCTATTGTGA	CTTTCTCTGA	CAATAAACAG	840
TTAGTCTTTG	AAAGAAACCA	TTCCATAATG	GGTGGCGGAG	CCATTTATGC	TAGGAAACTT	900
AGCATCTCTT	CAGGAGGTCC	TACTCTATTT	ATCAATAATA	TATCATATGC	AAATTCGCAA	960
AATTTAGGTG	GAGCTATTGC	CATTGATACT	GGAGGGGAGA	TCAGTTTATC	AGCAGAGAAA	1020
GGAACAATTA	CATTCCAAGG	AAACCGGACG	AGCTTACCGT	TTTTGAATGG	CATCCATCTT	1080
TTACAAAATG	CTAAATTCCT	GAAATTACAG	GCGAGAAATG	GATGCTCTAT	AGAATTTTAT	1140
GATCCTATTA	CTTCTGAAGC	AGATGGGTCT	ACCCAATTGA	ATATCAACGG	AGATCCTAAA	1200
AATAAAGAGT	ACACAGGGAC	CATACTCTTT	TCTGGAGAAA	AGAGTCTAGC	AAACGATCCT	1260
AGGGATTTTA	AATCTACAAT	CCCTCAGAAC	GTCAACCTGT	CTGCAGGATA	CTTAGTTATT	1320
AAAGAGGGGG	CCGAAGTCAC	AGTTTCAAAA	TTCACGCAGT	CTCCAGGATC	GCATTTAGTT	1380
TTAGATTTAG	GAACCAAATG	GATAGCCTCT	AAGGAAGACA	TTGCCATCAC	AGGCCTCGCG	1440
ATAGATATAG	ATAGCTTAAG	CTCATCCTCA	ACAGCAGCTG	TTATTAAAGC	AAACACCGCA	1500
AATAAACAGA	TATCCGTGAC	GGACTCTATA	GAACCTATCT	CGCCTACTGG	CAATGCCTAT	1560
GAAGATCTCA	GAATGAGAAA	TTCACAGACG	TTCCCTCTGC	TCTCTTTAGA	GCCTGGAGCC	1620
GGGGGTAGTG	TGACTGTAAC	TGCTGGAGAT	TTCTTACCGG	TAAGTCCCCA	TTATGGTTTT	1680
CAAGGCAATT	GGAAATTAGC	TTGGACAGGA	ACTGGAAACA	AAGTTGGAGA	ATTCTTCTGG	1740
GATAAAATAA	ATTATAAGCC	TAGACCTGAA	AAAGAAGGAA	ATTTAGTTCC	TAATATCTTG	1800
TGGGGGAATG	CTGTAAATGT	CAGATCCTTA	ATGCAGGTTT	AAGAGACCCA	TGCATCGAGC	1860
TTACAGACAG	ATCGAGGGCT	GTGGATCGAT	GGAATTGGGA	ATTTCTTCCA	TGTATCTGCC	1920
TCCGAAGACA	ATATAAGTCA	CCGTCAATAC	AGCGGTGGAT	ATGTTCTATC	TGTAAATAAT	1980
GAGATCACAC	CTAAGCACTA	TACTTCGATG	GCATTTTCCC	AACTCTTTAG	TAGAGACAAG	2040
GACTATGCGG	TTTCCAACAA	CGAATACAGA	ATGTATTTAG	GATCGTATCT	CTATCAATAT	2100
ACAACCTCCC	TAGGGAATAT	TTTCCGTTAT	GCTTCGCGTA	ACCCTAATGT	AAACGTCGGG	2160
ATTCTCTCAA	GAAGGTTTCT	TCAAAATCCT	CTTATGATTT	TTTATTTTTT	GTGTGCTTAT	2220
GGTCATGCCA	CCAATGATAT	GAAAACAGAC	TACGCAAATT	TCCCTATGGT	GAAAAACAGC	2280
TGGAGAAACA	ATTGTTGGGC	TATAGAGTGC	GGAGGGAGCA	TGCCTCTATT	GGTATTTGAG	2340
AACGGAAGAC	TTTTCCAAGG	TGCCATCCCA	TTTATGAAAC	TACAATTAGT	TTATGCTTAT	2400
CAGGGAGATT	TCAAAGAGAC	GACTGCAGAT	GGCCGTAGAT	TTAGTAATGG	GAGTTTAACA	2460
TCGATTTCTG	TACCTCTAGG	CATACGCTTT	GAGAAGCTGG	CACCTTCTCA	GGATGTACTC	2520
TATGACTTTA	GTTTCTCCTA	TATTCCTGAT	ATTTTCCGTA	AGGATCCCTC	ATGTGAAGCT	2580
GCTCTGGTGA	TTAGCGGAGA	CTCCTGGCTT	GTTCCGGCAG	CACACGTATC	AAGACATGCT	2640
TTTGTAGGGA	GTGGAACGGG	TCGGTATCAC	TTTAACGACT	ATACTGAGCT	CTTATGTCTA	2700
GGAAGTATAG	AATGCCGCCC	CCATGCTAGG	AATTATAATA	TAACTGTGG	AAGCAAATTT	2760
CGTTTTTTAGA	AGGTTTCCAT	TGCCTGTGTG	GTTCCGGATC	TTAACTATAA	ATCCTGGACT	2820
ATGGATCATA	GGCATTGGGT	TTCTCGAACT	TGTGTGGAGA	ATAACGACAT	TTTATATGCA	2880
TAACGGAATA	CTCGTATCAC	CTCAGCCCCC	AGAGACATTC	TTAGGGGTT	CTTTATTTGT	2940
CTAAACTTCG	TATTTTATCG	AGAATCCTTT	ACGTTCTTGG	TTTGCTTGTC	TCCGAGGAGT	3000
TCTCTAACGA	ATCATAGGGA	TTCCAGGGTT	CTGTTCTTGG	AGTCCTTTGG	A	3052

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 922 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Arg Phe Ser Leu Cys Gly Phe Pro Leu Val Phe Ser Leu Thr Leu
 1           5           10           15
Leu Ser Val Phe Asp Thr Ser Leu Ser Ala Thr Thr Ile Ser Leu Thr
          20           25           30
Pro Glu Asp Ser Phe His Gly Asp Ser Gln Asn Ala Glu Arg Ser Tyr
          35           40           45
Asn Val Gln Ala Gly Asp Val Tyr Ser Leu Thr Gly Asp Val Ser Ile
          50           55           60
Ser Asn Val Asp Asn Ser Ala Leu Asn Lys Ala Cys Phe Asn Val Thr
          65           70           75           80
Ser Gly Ser Val Thr Phe Ala Gly Asn His His Gly Leu Tyr Phe Asn
          85           90           95
Asn Ile Ser Ser Gly Thr Thr Lys Glu Gly Ala Val Leu Cys Cys Gln
          100          105          110
Asp Pro Gln Ala Thr Ala Arg Phe Ser Gly Phe Ser Thr Leu Ser Phe
          115          120          125
Ile Gln Ser Pro Gly Asp Ile Lys Glu Gln Gly Cys Leu Tyr Ser Lys
          130          135          140
Asn Ala Leu Met Leu Leu Asn Asn Tyr Val Val Arg Phe Glu Gln Asn
          145          150          155          160
Gln Ser Lys Thr Lys Gly Gly Ala Ile Ser Gly Ala Asn Val Thr Ile
          165          170          175
Val Gly Asn Tyr Asp Ser Val Ser Phe Tyr Gln Asn Ala Ala Thr Phe
          180          185          190
Gly Gly Ala Ile His Ser Ser Gly Pro Leu Gln Ile Ala Val Asn Gln
          195          200          205
Ala Glu Ile Arg Phe Ala Gln Asn Thr Ala Lys Asn Gly Ser Gly Gly
          210          215          220
Ala Leu Tyr Ser Asp Gly Asp Ile Asp Ile Asp Gln Asn Ala Tyr Val
          225          230          235          240
Leu Phe Arg Glu Asn Glu Ala Leu Thr Thr Ala Ile Gly Lys Gly Gly
          245          250          255
Ala Val Cys Cys Leu Pro Thr Ser Gly Ser Ser Thr Pro Val Pro Ile
          260          265          270

```

Val	Thr	Phe	Ser	Asp	Asn	Lys	Gln	Leu	Val	Phe	Glu	Arg	Asn	His	Ser		
		275					280					285					
Ile	Met	Gly	Gly	Gly	Ala	Ile	Tyr	Ala	Arg	Lys	Leu	Ser	Ile	Ser	Ser		
	290					295					300						
Gly	Gly	Pro	Thr	Leu	Phe	Ile	Asn	Asn	Ile	Ser	Tyr	Ala	Asn	Ser	Gln		
305					310					315					320		
Asn	Leu	Gly	Gly	Ala	Ile	Ala	Ile	Asp	Thr	Gly	Gly	Glu	Ile	Ser	Leu		
				325					330					335			
Ser	Ala	Glu	Lys	Gly	Thr	Ile	Thr	Phe	Gln	Gly	Asn	Arg	Thr	Ser	Leu		
			340					345					350				
Pro	Phe	Leu	Asn	Gly	Ile	His	Leu	Leu	Gln	Asn	Ala	Lys	Phe	Leu	Lys		
		355					360					365					
Leu	Gln	Ala	Arg	Asn	Gly	Cys	Ser	Ile	Glu	Phe	Tyr	Asp	Pro	Ile	Thr		
	370					375					380						
Ser	Glu	Ala	Asp	Gly	Ser	Thr	Gln	Leu	Asn	Ile	Asn	Gly	Asp	Pro	Lys		
385					390					395					400		
Asn	Lys	Glu	Tyr	Thr	Gly	Thr	Ile	Leu	Phe	Ser	Gly	Glu	Lys	Ser	Leu		
				405					410					415			
Ala	Asn	Asp	Pro	Arg	Asp	Phe	Lys	Ser	Thr	Ile	Pro	Gln	Asn	Val	Asn		
			420					425					430				
Leu	Ser	Ala	Gly	Tyr	Leu	Val	Ile	Lys	Glu	Gly	Ala	Glu	Val	Thr	Val		
		435					440					445					
Ser	Lys	Phe	Thr	Gln	Ser	Pro	Gly	Ser	His	Leu	Val	Leu	Asp	Leu	Gly		
	450					455					460						
Thr	Lys	Leu	Ile	Ala	Ser	Lys	Glu	Asp	Ile	Ala	Ile	Thr	Gly	Leu	Ala		
465					470					475					480		
Ile	Asp	Ile	Asp	Ser	Leu	Ser	Ser	Ser	Ser	Thr	Ala	Ala	Val	Ile	Lys		
				485					490					495			
Ala	Asn	Thr	Ala	Asn	Lys	Gln	Ile	Ser	Val	Thr	Asp	Ser	Ile	Glu	Leu		
			500					505					510				
Ile	Ser	Pro	Thr	Gly	Asn	Ala	Tyr	Glu	Asp	Leu	Arg	Met	Arg	Asn	Ser		
		515					520					525					
Gln	Thr	Phe	Pro	Leu	Leu	Ser	Leu	Glu	Pro	Gly	Ala	Gly	Gly	Ser	Val		
	530					535					540						
Thr	Val	Thr	Ala	Gly	Asp	Phe	Leu	Pro	Val	Ser	Pro	His	Tyr	Gly	Phe		
545					550					555					560		
Gln	Gly	Asn	Trp	Lys	Leu	Ala	Trp	Thr	Gly	Thr	Gly	Asn	Lys	Val	Gly		
				565					570					575			
Glu	Phe	Phe	Trp	Asp	Lys	Ile	Asn	Tyr	Lys	Pro	Arg	Pro	Glu	Lys	Glu		
			580					585					590				



Gly	Asn	Leu	Val	Pro	Asn	Ile	Leu	Trp	Gly	Asn	Ala	Val	Asn	Val	Arg	595	600	605
Ser	Leu	Met	Gln	Val	Gln	Glu	Thr	His	Ala	Ser	Ser	Leu	Gln	Thr	Asp	610	615	620
Arg	Gly	Leu	Trp	Ile	Asp	Gly	Ile	Gly	Asn	Phe	Phe	His	Val	Ser	Ala	625	630	635
Ser	Glu	Asp	Asn	Ile	Arg	Tyr	Arg	His	Asn	Ser	Gly	Gly	Tyr	Val	Leu	645	650	655
Ser	Val	Asn	Asn	Glu	Ile	Thr	Pro	Lys	His	Tyr	Thr	Ser	Met	Ala	Phe	660	665	670
Ser	Gln	Leu	Phe	Ser	Arg	Asp	Lys	Asp	Tyr	Ala	Val	Ser	Asn	Asn	Glu	675	680	685
Tyr	Arg	Met	Tyr	Leu	Gly	Ser	Tyr	Leu	Tyr	Gln	Tyr	Thr	Thr	Ser	Leu	690	695	700
Gly	Asn	Ile	Phe	Arg	Tyr	Ala	Ser	Arg	Asn	Pro	Asn	Val	Asn	Val	Gly	705	710	715
Ile	Leu	Ser	Arg	Arg	Phe	Leu	Gln	Asn	Pro	Leu	Met	Ile	Phe	His	Phe	725	730	735
Leu	Cys	Ala	Tyr	Gly	His	Ala	Thr	Asn	Asp	Met	Lys	Thr	Asp	Tyr	Ala	740	745	750
Asn	Phe	Pro	Met	Val	Lys	Asn	Ser	Trp	Arg	Asn	Asn	Cys	Trp	Ala	Ile	755	760	765
Glu	Cys	Gly	Gly	Ser	Met	Pro	Leu	Leu	Val	Phe	Glu	Asn	Gly	Arg	Leu	770	775	780
Phe	Gln	Gly	Ala	Ile	Pro	Phe	Met	Lys	Leu	Gln	Leu	Val	Tyr	Ala	Tyr	785	790	795
Gln	Gly	Asp	Phe	Lys	Glu	Thr	Thr	Ala	Asp	Gly	Arg	Arg	Phe	Ser	Asn	805	810	815
Gly	Ser	Leu	Thr	Ser	Ile	Ser	Val	Pro	Leu	Gly	Ile	Arg	Phe	Glu	Lys	820	825	830
Leu	Ala	Leu	Ser	Gln	Asp	Val	Leu	Tyr	Asp	Phe	Ser	Phe	Ser	Tyr	Ile	835	840	845
Pro	Asp	Ile	Phe	Arg	Lys	Asp	Pro	Ser	Cys	Glu	Ala	Ala	Leu	Val	Ile	850	855	860
Ser	Gly	Asp	Ser	Trp	Leu	Val	Pro	Ala	Ala	His	Val	Ser	Arg	His	Ala	865	870	875
Phe	Val	Gly	Ser	Gly	Thr	Gly	Arg	Tyr	His	Phe	Asn	Asp	Tyr	Thr	Glu	885	890	895

Asn Ile Asn Cys Gly Ser Lys Phe Arg Phe  
915 920

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 2526 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGAAGATTCC	CACTCCGCTT	TTTATTGATA	TCATTAGTAC	CTACGCTTTC	TATGTCGAAT	60
TTATTAGGAG	CTGCTACTAC	CGAAGAGCTA	TCGGCTAGCA	ATAGCTTCGA	TGGAAC TACA	120
TCAACAACAA	GCTTTTCTAG	TAAAACATCA	TCGGCTACAG	ATGGCACCAA	TTATGTTTTT	180
AAAGATTCTG	TAGTTATAGA	AAATGTACCC	AAAACAGGGG	AAACTCAGTC	TACTAGTTGT	240
TTTAAAAATG	ACGCTGCAGC	TGGAGATCTA	AATTTCTTAG	GAGGGGGATT	TTCTTTCACA	300
TTTAGCAATA	TCGATGCAAC	CACGGCTTCT	GGAGCTGCTA	TTGGAAGTGA	AGCAAGCTAAT	360
AAGACAGTCA	CGTTATCAGG	ATTTTCGGCA	CTTTCTTTTC	TTAAATCCCC	AGCAAGTACA	420
GTGACTAATG	GATTGGGAGC	TATCAATGTT	AAAGGGAATT	TAAGCCTATT	GGATAATGAT	480
AAGGTATTGA	TTCAGGACAA	TTTCTCAACA	GGAGATGGCG	GAGCAATTAA	TTGTGCAGGC	540
TCCTTGAAGA	TCGCAAACAA	TAAGTCCCTT	TCTTTTATTG	GAAATAGTTC	TTCAACACGT	600
GGCGGAGCGA	TTCATACCAA	AAACCTCACA	CTATCTTCTG	GTGGGGAAAC	TCTATTTT CAG	660
GGGAATACAG	CGCCTACGGC	TGCTGGTAAA	GGAGGTGCTA	TCGCGATTGC	AGACTCTGGC	720
ACCCTATCCA	TTTCTGGAGA	CAGTGGCGCAG	ATTATCTTTG	AAGGCAATAC	GATAGGAGCT	780
ACAGGAACCG	TCTCTCATAG	TGCTATTGAT	TTAGGACTTA	CGCCTAAGAT	AACTGCGTTA	840
CGTGCTGCGC	AAGGACATAC	GATATACTTT	TATGATCCGA	TTACTGTAAC	AGGATCGACA	900
TCTGTTGCTG	ATGCTCTCAA	TATTAATAGC	CCTGATACTG	GAGATAACAA	AGAGTATACG	960
GGAACCATAG	TCTTTTCTGG	AGAGAAGCTC	ACGGAGGCAG	AAGCTAAAGA	TGAGAAGAAC	1020
CGCACTTCTA	AATTACTTCA	AAATGTTGCT	TTTAAAAATG	GGACTGTAGT	TTTAAAAGGT	1080
GATGTCGTTT	TAAGTGCGAA	CGTTTCTCT	CAGGATGCAA	ACTCTAAGTT	GATTATGGAT	1140
TTAGGGACGT	CGTTGGTTGC	AAACACCGAA	AGTATCGAGT	TAACGAATTT	GGAAATTAAT	1200
ATAGACTCTC	TCAGGAACGG	GAAAAGATA	AAACTCAGTG	CTGCCACAGC	TCAGAAAGAT	1260
ATTTCGTATAG	ATCGTCTGT	TGTACTGGCA	ATTAGCGATG	AGAGTTTTTTA	TCAAAATGGC	1320
TTTTTGAAATG	AGGACCATT C	CTATGATGGG	ATTCTTGAGT	TAGATGCTGG	GAAAGACATC	1380
GTGATTTCTG	CAGATTCTCG	CAGTATAAAT	GCTGTACAAT	CTCCGTATGG	CTATCAGGGA	1440
AAGTGGACAA	TCAATTGGTC	TACTGATGAT	AAGAAAGCTA	CGGTTTCTTG	GGCAAAGCAA	1500
AGTTTTAATC	CCACTGCTGA	GCAGGAGGCT	CCGTTAGTTC	CTAATCTTCT	TTGGGGTTCT	1560
TTTATAGATG	TTCGTCCCTT	CCAAAATTTT	ATAGAGCTAG	GTACTGAAGG	TGCTCCTTAC	1620
GAAAAGAGAT	TTTGGGTTGC	AGGCATTTCC	AATGTTTTTG	ATAGGAGCGG	TCGTGAAAAT	1680
CAAAGGAAAT	TCCGT CATGT	GAGTGGAGGT	GCTGTAGTAG	GTGCTAGCAC	GAGGATGCCG	1740
GGTGGTGATA	CCTTGCTCTCT	GGGTTTTGCT	CAGCTCTTTG	CGCTTGACAA	AGACTACTTT	1800
ATGAATACCA	ATTTTCGAAA	GACCTACGCA	GGATCTTTAC	GGTTGCGACA	CGATGCTTCC	1860
CTATACTCTG	TGGTGAGTAT	CCTTTTAGGA	GAGGGAGGAC	TCCGCGAGAT	CCTGTTGCCT	1920
TATGTTTCCA	AGACTCTGCC	GTGCTCTTTC	TATGGGCAGC	TTAGCTACGG	CCATACGGAT	1980
CATCGCATGA	AGACCGAGTC	TCTACCCCCC	CCCCCCCCGA	CGCTCTCGAC	GGATCATACT	2040
TCTTGGGGAG	GATATGTCTG	GGCTGGAGAG	CTGGGAATC	GAGTTGCTGT	TGAAAATACC	2100
AGCGGCAGAG	GATTTTTTCCG	AGAGTACACT	CCATTTGTAA	AAGTCCAAGC	TGTTTACTCG	2160
CGCCAAGATA	GCTTTGTTCT	ACTAGGAGCT	ATCAGTCGTG	ATTTTAGTGA	TTCGCATCTT	2220
TATAACCTTG	CGATTCTCTG	TGGAATCAAG	TTAGAGAAAC	GGTTTG CAGA	GCAATATTAT	2280
CATGTTGTAG	CGATGTATT C	TCCAGATGTT	TGTCGTAGTA	ACCCCAAATG	TACGACTACC	2340
CTACTTTCCA	ACCAAGGGAG	TTGGAAGACC	AAAGGTTCTGA	ACTTAGCAAG	ACAGGCTGGT	2400

```

ATTGTTTCAGG CCTCAGGTTT TCGATCTTTG GGAGCTGCAG CAGAGCTTTT CGGGAACTTT 2460
GGCTTTGAAT GGCGGGGATC TTCTCGTAGC TATAATGTAG ATGCGGGTAG CAAAATCAAA 2520
TTTTAG 2526

```

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 841 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Lys Ile Pro Leu Arg Phe Leu Leu Ile Ser Leu Val Pro Thr Leu
 1             5             10             15
Ser Met Ser Asn Leu Leu Gly Ala Ala Thr Thr Glu Glu Leu Ser Ala
          20             25             30
Ser Asn Ser Phe Asp Gly Thr Thr Ser Thr Thr Ser Phe Ser Ser Lys
          35             40             45
Thr Ser Ser Ala Thr Asp Gly Thr Asn Tyr Val Phe Lys Asp Ser Val
          50             55             60
Val Ile Glu Asn Val Pro Lys Thr Gly Glu Thr Gln Ser Thr Ser Cys
          65             70             75             80
Phe Lys Asn Asp Ala Ala Ala Gly Asp Leu Asn Phe Leu Gly Gly Gly
          85             90             95
Phe Ser Phe Thr Phe Ser Asn Ile Asp Ala Thr Thr Ala Ser Gly Ala
          100            105            110
Ala Ile Gly Ser Glu Ala Ala Asn Lys Thr Val Thr Leu Ser Gly Phe
          115            120            125
Ser Ala Leu Ser Phe Leu Lys Ser Pro Ala Ser Thr Val Thr Asn Gly
          130            135            140
Leu Gly Ala Ile Asn Val Lys Gly Asn Leu Ser Leu Leu Asp Asn Asp
          145            150            155            160
Lys Val Leu Ile Gln Asp Asn Phe Ser Thr Gly Asp Gly Gly Ala Ile
          165            170            175
Asn Cys Ala Gly Ser Leu Lys Ile Ala Asn Asn Lys Ser Leu Ser Phe
          180            185            190
Ile Gly Asn Ser Ser Ser Thr Arg Gly Gly Ala Ile His Thr Lys Asn
          195            200            205
Leu Thr Leu Ser Ser Gly Gly Glu Thr Leu Phe Gln Gly Asn Thr Ala
          210            215            220

```

Pro Thr Ala Ala Gly Lys Gly Gly Ala Ile Ala Ile Ala Asp Ser Gly  
 225 230 235 240  
 Thr Leu Ser Ile Ser Gly Asp Ser Gly Asp Ile Ile Phe Glu Gly Asn  
 245 250 255  
 Thr Ile Gly Ala Thr Gly Thr Val Ser His Ser Ala Ile Asp Leu Gly  
 260 265 270  
 Thr Ser Ala Lys Ile Thr Ala Leu Arg Ala Ala Gln Gly His Thr Ile  
 275 280 285  
 Tyr Phe Tyr Asp Pro Ile Thr Val Thr Gly Ser Thr Ser Val Ala Asp  
 290 295 300  
 Ala Leu Asn Ile Asn Ser Pro Asp Thr Gly Asp Asn Lys Glu Tyr Thr  
 305 310 315 320  
 Gly Thr Ile Val Phe Ser Gly Glu Lys Leu Thr Glu Ala Glu Ala Lys  
 325 330 335  
 Asp Glu Lys Asn Arg Thr Ser Lys Leu Leu Gln Asn Val Ala Phe Lys  
 340 345 350  
 Asn Gly Thr Val Val Leu Lys Gly Asp Val Val Leu Ser Ala Asn Gly  
 355 360 365  
 Phe Ser Gln Asp Ala Asn Ser Lys Leu Ile Met Asp Leu Gly Thr Ser  
 370 375 380  
 Leu Val Ala Asn Thr Glu Ser Ile Glu Leu Thr Asn Leu Glu Ile Asn  
 385 390 395 400  
 Ile Asp Ser Leu Arg Asn Gly Lys Lys Ile Lys Leu Ser Ala Ala Thr  
 405 410 415  
 Ala Gln Lys Asp Ile Arg Ile Asp Arg Pro Val Val Leu Ala Ile Ser  
 420 425 430  
 Asp Glu Ser Phe Tyr Gln Asn Gly Phe Leu Asn Glu Asp His Ser Tyr  
 435 440 445  
 Asp Gly Ile Leu Glu Leu Asp Ala Gly Lys Asp Ile Val Ile Ser Ala  
 450 455 460  
 Asp Ser Arg Ser Ile Asn Ala Val Gln Ser Pro Tyr Gly Tyr Gln Gly  
 465 470 475 480  
 Lys Trp Thr Ile Asn Trp Ser Thr Asp Asp Lys Lys Ala Thr Val Ser  
 485 490 495  
 Trp Ala Lys Gln Ser Phe Asn Pro Thr Ala Glu Gln Glu Ala Pro Leu  
 500 505 510  
 Val Pro Asn Leu Leu Trp Gly Ser Phe Ile Asp Val Arg Pro Phe Gln  
 515 520 525  
 Asn Phe Ile Glu Leu Gly Thr Glu Gly Ala Pro Tyr Glu Lys Arg Phe

530	535	540
Trp Val Ala Gly Ile Ser Asn Val Leu His Arg Ser Gly Arg Glu Asn 545 550 555 560		
Gln Arg Lys Phe Arg His Val Ser Gly Gly Ala Val Val Gly Ala Ser 565 570 575		
Thr Arg Met Pro Gly Gly Asp Thr Leu Ser Leu Gly Phe Ala Gln Leu 580 585 590		
Phe Ala Arg Asp Lys Asp Tyr Phe Met Asn Thr Asn Phe Ala Lys Thr 595 600 605		
Tyr Ala Gly Ser Leu Arg Leu Gln His Asp Ala Ser Leu Tyr Ser Val 610 615 620		
Val Ser Ile Leu Leu Gly Glu Gly Gly Leu Arg Glu Ile Leu Leu Pro 625 630 635 640		
Tyr Val Ser Lys Thr Leu Pro Cys Ser Phe Tyr Gly Gln Leu Ser Tyr 645 650 655		
Gly His Thr Asp His Arg Met Lys Thr Glu Ser Leu Pro Pro Pro Pro 660 665 670		
Pro Thr Leu Ser Thr Asp His Thr Ser Trp Gly Gly Tyr Val Trp Ala 675 680 685		
Gly Glu Leu Gly Thr Arg Val Ala Val Glu Asn Thr Ser Gly Arg Gly 690 695 700		
Phe Phe Arg Glu Tyr Thr Pro Phe Val Lys Val Gln Ala Val Tyr Ser 705 710 715 720		
Arg Gln Asp Ser Phe Val Glu Leu Gly Ala Ile Ser Arg Asp Phe Ser 725 730 735		
Asp Ser His Leu Tyr Asn Leu Ala Ile Pro Leu Gly Ile Lys Leu Glu 740 745 750		
Lys Arg Phe Ala Glu Gln Tyr Tyr His Val Val Ala Met Tyr Ser Pro 755 760 765		
Asp Val Cys Arg Ser Asn Pro Lys Cys Thr Thr Thr Leu Leu Ser Asn 770 775 780		
Gln Gly Ser Trp Lys Thr Lys Gly Ser Asn Leu Ala Arg Gln Ala Gly 785 790 795 800		
Ile Val Gln Ala Ser Gly Phe Arg Ser Leu Gly Ala Ala Ala Glu Leu 805 810 815		
Phe Gly Asn Phe Gly Phe Glu Trp Arg Gly Ser Ser Arg Ser Tyr Asn 820 825 830		
Val Asp Ala Gly Ser Lys Ile Lys Phe 835 840		

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2787 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATGAAGTCTT	CTTTCCCCAA	GTTTGTATTT	TCTACATTTG	CTATTTTCCC	TTTGTCTATG	60
ATTGCTACCG	AGACAGTTTT	GGATTCAAGT	GCGAGTTTCG	ATGGGAATAA	AAATGGTAAT	120
TTTTCAAGTT	GTGAGAGTCA	GGAAGATGCT	GGAACCTACCT	ACCTATTTAA	GGGAAATGTC	180
ACTCTAGAAA	ATATTCCCTGG	AACAGGCACA	GCAATCACAA	AAAGCTGTTT	TAACAACACT	240
AAGGGCGATT	TGACTTTTAC	AGGTAACGGG	AACCTCTCTAT	TGTTCCAAAC	GGTGGATGCA	300
GGGACTGTAG	CAGGGGCTGC	TGTTAACAGC	AGCGTGGTAG	ATAAATCTAC	CACGTTTATA	360
GGGTTTTCTT	CGCTATCTTT	TATTGCGTCT	CCTGGAAGTT	CGATAACTAC	CGGCAAAGGA	420
GCCGTTAGCT	GCTCTACGGG	TAGCTTGAA	TTTGACAAA	ATGTCAGTTT	GCTCTTCAGC	480
AAAACTTTT	CAACGGATAA	TGGCGGTGCT	ATCACCGCAA	AACTCTTTC	ATTAACAGGG	540
ACTACAATGT	CAGCTCTGTT	TTCTGAAAAT	ACCTCCTCAA	AGAAAGGCGG	AGCCATTCAG	600
ACTTCCGATG	CCCTTACCAT	TACTGGAAAC	CAAGGGGAAG	TCTCTTTTTC	TGACAATACT	660
TCTTCGGATT	CTGGAGCTGC	AATTTTACA	GAAGCCTCGG	TGACTATTTT	TAATAATGCT	720
AAAGTTTCCT	TTATTGACAA	TAAGGTCACA	GGAGCGAGCT	CCTCAACAAC	GGGGGATATG	780
TCAGGAGGTG	CTATCTGTGC	TTATAAAACT	AGTACAGATA	CTAAGGTCAC	CCTCACTGGA	840
AATCAGATGT	TACTCTTCAG	CAACAATACA	TCGACAACAG	CGGGAGGAGC	TATCTATGTG	900
AAAAAGCTCG	AACTGGCTTC	CGGAGGACTT	ACCCTATTCA	GTAGAAATAG	TGTCAATGGA	960
GGTACAGCTC	CTAAAGGTGG	AGCCATAGCT	ATCGAAGATA	GTGGGGAATT	GAGTTTATCC	1020
GCCGATAGTG	GTGACATTGT	CTTTTTAGGG	AATACAGTCA	CTTCTACTAC	TCCTGGGACG	1080
AATAGAAGTA	GTATCGACTT	AGGAACGAGT	GCAAAGATGA	CAGCTTTGCG	TTCTGCTGCT	1140
GGTAGAGCCA	TCTACTTCTA	TGATCCCAT	ACTACAGGAT	CTTCCACAAC	AGTTACAGAT	1200
GTCTTAAAG	TTAATGAGAC	TCCGGCAGAT	TCTGCACACT	AATATACAGG	GAACATCATC	1260
TTCACAGGAG	AAAAGTTATC	AGAGACAGAG	GCCGCAGATT	CTAAAAATCT	TACTTCGAAG	1320
CTACTACAGC	CTGTAACCTC	TTCAGGAGGT	ACTCTATCTT	TAAAACATGG	AGTGACTCTG	1380
CAGACTCAGG	CATTCACTCA	ACAGGCAGAT	TCTCGTCTCG	AAATGGACGT	AGGAACTACT	1440
CTAGAACCTG	CTGATACTAG	CACCATAAAC	AATTTGGTCA	TTAACATCAG	TTCTATAGAC	1500
GGTGCAAAGA	AGGCAAAAAT	AGAAACCAAA	GCTACGTCAA	AAAATCTGAC	TTTATCTGGA	1560
ACCATCACTT	TATTGGACCC	GACGGGCACG	TTTTATGAAA	ATCATAGTTT	AAGAAATCCT	1620
CAGTCCTACG	ACATCTTAGA	GCTCAAAGCT	TCTGGAAGTG	TAACAAGCAC	CGCAGTGACT	1680
CCAGATCCTA	TAATGGGTGA	GAAATTCAT	TACGGCTATC	AGGGAACTTG	GGGCCCAATT	1740
GTTTGGGGGA	CAGGGGCTTC	TACGACTGCA	ACCTTCAACT	GGACTAAAAC	TGGCTATATT	1800
CCTAATCCCG	AGCGTATCGG	CTCTTTAGTC	CCTAATAGCT	TATGGAATGC	ATTTATAGAT	1860
ATTAGCTCTC	TCCATTATCT	TATGGAGACT	GCAAACGAAG	GGTTGCAGGG	AGACCGTGCT	1920
TTTTGGTGTG	CTGGATTATC	TAACCTCTTC	CATAAGGATA	GTACAAAAAC	ACGACGCGGG	1980
TTTCGCCATT	TGAGTGGCGG	TTATGTCATA	GGAGGAAACC	TACATACTTG	TTCAGATAAG	2040
ATTCTTAGTG	CTGCATTTTG	TCAGCTCTTT	GGAAGAGATA	GAGACTACTT	TGTAGCTAAG	2100
AATCAAGGTA	CAGTCTACGG	AGGAACTCTC	TATTACCAGC	ACAACGAAAC	CTATATCTCT	2160
CTTCCTTGCA	AACTACGGCC	TTGTTTCGTT	TCTTATGTTT	CTACAGAGAT	TCCTGTTCTC	2220
TTTTCAGGAA	ACCTTAGCTA	CACCCATACG	GATAACGATC	TGAAAACCAA	GTATACAACA	2280
TATCCTACTG	TTAAAGGAAG	CTGGGGGAAT	GATAGTTTCG	CTTTAGAATT	CGGTGGAAGA	2340
GCTCCGATTT	GCTTAGATGA	AAGTGCTCTA	TTTGAGCAGT	ACATGCCCTT	CATGAAATTG	2400
CAGTTTGTCT	ATGCACATCA	GGAAGGTTTT	AAAGAACAGG	GAACAGAAGC	TCGTGAATTT	2460
GGAAGTAGCC	GTCTTGTGAA	TCTTGCCTTA	CCTATCGGGA	TCCGATTTGA	TAAGGAATCA	2520
GACTGCCAAG	ATGCAACGTA	CAATCTAACT	CTTGGTTATA	CTGTGGATCT	TGTTTCGTAGT	2580
AACCCCGACT	GTACGACAAC	ACTGCGAATT	AGCGGTGATT	CTTGGAAAAC	CTTCGGTACG	2640
AATTTGGCAA	GACAAGCTTT	AGTCCTTCGT	GCAGGGAACC	ATTTTGTGCTT	TAACCTCAAAT	2700

TTTGAAGCCT TTAGCCAATT TTCTTTTGAA TTGCGTGGGT CATCTCGCAA TTACAATGTA 2760  
GACTTAGGAG CAAAATACCA ATTCTAA 2787

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 928 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met 1	Lys	Ser	Ser	Phe 5	Pro	Lys	Phe	Val	Phe 10	Ser	Thr	Phe	Ala	Ile 15	Phe
Pro	Leu	Ser	Met 20	Ile	Ala	Thr	Glu	Thr 25	Val	Leu	Asp	Ser	Ser 30	Ala	Ser
Phe	Asp	Gly 35	Asn	Lys	Asn	Gly	Asn 40	Phe	Ser	Val	Arg	Glu 45	Ser	Gln	Glu
Asp	Ala 50	Gly	Thr	Thr	Tyr	Leu 55	Phe	Lys	Gly	Asn	Val 60	Thr	Leu	Glu	Asn
Ile 65	Pro	Gly	Thr	Gly	Thr 70	Ala	Ile	Thr	Lys	Ser 75	Cys	Phe	Asn	Asn	Thr 80
Lys	Gly	Asp	Leu	Thr 85	Phe	Thr	Gly	Asn	Gly 90	Asn	Ser	Leu	Leu	Phe 95	Gln
Thr	Val	Asp	Ala 100	Gly	Thr	Val	Ala	Gly	Ala 105	Ala	Val	Asn	Ser	Ser	Val
Val	Asp	Lys 115	Ser	Thr	Thr	Phe	Ile 120	Gly	Phe	Ser	Ser 125	Leu	Ser	Phe	Ile
Ala	Ser 130	Pro	Gly	Ser	Ser	Ile 135	Thr	Thr	Gly	Lys	Gly 140	Ala	Val	Ser	Cys
Ser 145	Thr	Gly	Ser	Leu	Lys 150	Phe	Asp	Lys	Asn	Val 155	Ser	Leu	Leu	Phe	Ser 160
Lys	Asn	Phe	Ser	Thr 165	Asp	Asn	Gly	Gly	Ala 170	Ile	Thr	Ala	Lys	Thr 175	Leu
Ser	Leu	Thr	Gly 180	Thr	Thr	Met	Ser	Ala 185	Leu	Phe	Ser	Glu	Asn 190	Thr	Ser
Ser	Lys	Lys 195	Gly	Gly	Ala	Ile	Gln 200	Thr	Ser	Asp	Ala	Leu	Thr 205	Ile	Thr
Gly 210	Asn	Gln	Gly	Glu	Val	Ser 215	Phe	Ser	Asp	Asn	Thr 220	Ser	Ser	Asp	Ser
Gly	Ala	Ala	Ile	Phe	Thr	Glu	Ala	Ser	Val	Thr	Ile	Ser	Asn	Asn	Ala

225		230		235		240
Lys Val Ser Phe Ile Asp Asn Lys Val Thr Gly Ala Ser Ser Ser Thr						
	245			250		255
Thr Gly Asp Met Ser Gly Gly Ala Ile Cys Ala Tyr Lys Thr Ser Thr						
	260		265			270
Asp Thr Lys Val Thr Leu Thr Gly Asn Gln Met Leu Leu Phe Ser Asn						
	275		280			285
Asn Thr Ser Thr Thr Ala Gly Gly Ala Ile Tyr Val Lys Lys Leu Glu						
	290		295			300
Leu Ala Ser Gly Gly Leu Thr Leu Phe Ser Arg Asn Ser Val Asn Gly						
305		310		315		320
Gly Thr Ala Pro Lys Gly Gly Ala Ile Ala Ile Glu Asp Ser Gly Glu						
	325		330			335
Leu Ser Leu Ser Ala Asp Ser Gly Asp Ile Val Phe Leu Gly Asn Thr						
	340		345			350
Val Thr Ser Thr Thr Pro Gly Thr Asn Arg Ser Ser Ile Asp Leu Gly						
	355		360			365
Thr Ser Ala Lys Met Thr Ala Leu Arg Ser Ala Ala Gly Arg Ala Ile						
	370		375			380
Tyr Phe Tyr Asp Pro Ile Thr Thr Gly Ser Ser Thr Thr Val Thr Asp						
385		390		395		400
Val Leu Lys Val Asn Glu Thr Pro Ala Asp Ser Ala Leu Gln Tyr Thr						
	405		410			415
Gly Asn Ile Ile Phe Thr Gly Glu Lys Leu Ser Glu Thr Glu Ala Ala						
	420		425			430
Asp Ser Lys Asn Leu Thr Ser Lys Leu Leu Gln Pro Val Thr Leu Ser						
	435		440			445
Gly Gly Thr Leu Ser Leu Lys His Gly Val Thr Leu Gln Thr Gln Ala						
	450		455			460
Phe Thr Gln Gln Ala Asp Ser Arg Leu Glu Met Asp Val Gly Thr Thr						
465		470		475		480
Leu Glu Pro Ala Asp Thr Ser Thr Ile Asn Asn Leu Val Ile Asn Ile						
	485		490			495
Ser Ser Ile Asp Gly Ala Lys Lys Ala Lys Ile Glu Thr Lys Ala Thr						
	500		505			510
Ser Lys Asn Leu Thr Leu Ser Gly Thr Ile Thr Leu Leu Asp Pro Thr						
	515		520			525
Gly Thr Phe Tyr Glu Asn His Ser Leu Arg Asn Pro Gln Ser Tyr Asp						
	530		535			540



Ile	Leu	Glu	Leu	Lys	Ala	Ser	Gly	Thr	Val	Thr	Ser	Thr	Ala	Val	Thr	545	550	555	560
Pro	Asp	Pro	Ile	Met	Gly	Glu	Lys	Phe	His	Tyr	Gly	Tyr	Gln	Gly	Thr	565	570	575	
Trp	Gly	Pro	Ile	Val	Trp	Gly	Thr	Gly	Ala	Ser	Thr	Thr	Ala	Thr	Phe	580	585	590	
Asn	Trp	Thr	Lys	Thr	Gly	Tyr	Ile	Pro	Asn	Pro	Glu	Arg	Ile	Gly	Ser	595	600	605	
Leu	Val	Pro	Asn	Ser	Leu	Trp	Asn	Ala	Phe	Ile	Asp	Ile	Ser	Ser	Leu	610	615	620	
His	Tyr	Leu	Met	Glu	Thr	Ala	Asn	Glu	Gly	Leu	Gln	Gly	Asp	Arg	Ala	625	630	635	640
Phe	Trp	Cys	Ala	Gly	Leu	Ser	Asn	Phe	Phe	His	Lys	Asp	Ser	Thr	Lys	645	650	655	
Thr	Arg	Arg	Gly	Phe	Arg	His	Leu	Ser	Gly	Gly	Tyr	Val	Ile	Gly	Gly	660	665	670	
Asn	Leu	His	Thr	Cys	Ser	Asp	Lys	Ile	Leu	Ser	Ala	Ala	Phe	Cys	Gln	675	680	685	
Leu	Phe	Gly	Arg	Asp	Arg	Asp	Tyr	Phe	Val	Ala	Lys	Asn	Gln	Gly	Thr	690	695	700	
Val	Tyr	Gly	Gly	Thr	Leu	Tyr	Tyr	Gln	His	Asn	Glu	Thr	Tyr	Ile	Ser	705	710	715	720
Leu	Pro	Cys	Lys	Leu	Arg	Pro	Cys	Ser	Leu	Ser	Tyr	Val	Pro	Thr	Glu	725	730	735	
Ile	Pro	Val	Leu	Phe	Ser	Gly	Asn	Leu	Ser	Tyr	Thr	His	Thr	Asp	Asn	740	745	750	
Asp	Leu	Lys	Thr	Lys	Tyr	Thr	Thr	Tyr	Pro	Thr	Val	Lys	Gly	Ser	Trp	755	760	765	
Gly	Asn	Asp	Ser	Phe	Ala	Leu	Glu	Phe	Gly	Gly	Arg	Ala	Pro	Ile	Cys	770	775	780	
Leu	Asp	Glu	Ser	Ala	Leu	Phe	Glu	Gln	Tyr	Met	Pro	Phe	Met	Lys	Leu	785	790	795	800
Gln	Phe	Val	Tyr	Ala	His	Gln	Glu	Gly	Phe	Lys	Glu	Gln	Gly	Thr	Glu	805	810	815	
Ala	Arg	Glu	Phe	Gly	Ser	Ser	Arg	Leu	Val	Asn	Leu	Ala	Leu	Pro	Ile	820	825	830	
Gly	Ile	Arg	Phe	Asp	Lys	Glu	Ser	Asp	Cys	Gln	Asp	Ala	Thr	Tyr	Asn	835	840	845	

Leu Thr Leu Gly Tyr Thr Val Asp Leu Val Arg Ser Asn Pro Asp Cys  
 850 855 860  
 Thr Thr Thr Leu Arg Ile Ser Gly Asp Ser Trp Lys Thr Phe Gly Thr  
 865 870 875 880  
 Asn Leu Ala Arg Gln Ala Leu Val Leu Arg Ala Gly Asn His Phe Cys  
 885 890 895  
 Phe Asn Ser Asn Phe Glu Ala Phe Ser Gln Phe Ser Phe Glu Leu Arg  
 900 905 910  
 Gly Ser Ser Arg Asn Tyr Asn Val Asp Leu Gly Ala Lys Tyr Gln Phe  
 915 920 925

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2757 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATGAGATCGT	CTTTTTCTT	GTTATTAATA	TCTTCATCTC	TAGCCTTTCC	TCTCTTAATG	60
AGTGTTCCTG	CAGATGCTGC	CGATCTCACA	TTAGGGAGTC	GTGACAGTTA	TAATGGTGAT	120
ACAAGCACCA	CAGAATTTAC	TCCTAAAGCG	GCAACTTCTG	ATGCTAGTGG	CACGACCTAT	180
ATTCTCGATG	GGGATGTCTC	GATAAGCCAA	GCAGGGAAAC	AAACGAGCTT	AACCACAACT	240
TGTTTTTCTA	ACACTGCAGG	AAATCTTACC	TTCTTAGGGA	ACGGATTTTC	TCTTCATTTT	300
GACAAATATTA	TTTCGTCTAC	TGTTGCAGGT	GTTGTTGTTA	GCAATACAGC	AGCTTCTGGG	360
ATTACGAAAT	TCTCAGGATT	TTCAACTCTT	CGATGCTTG	CAGCTCCTAG	GACCACAGGT	420
AAAGGAGCCA	TTAAAATTAC	CGATGGTCTG	GTGTTTGAGA	GTATAGGGAA	TCTTGACCAA	480
AATGAAAATG	CCTCTAGTGA	AAATGGGGGA	GCCATCAATA	CGAAGACTTT	GTCTTTGACT	540
GGGAGTACGC	GGTTTGTAGC	GTTCTTGGC	AATAGCTCGT	CGCAACAAGG	GGGAGCGATC	600
TATGCTTCTG	GTGACTCTGT	GATTTCTGAG	AATGCAGGAA	TCTTGAGCTT	CGGAAACAAC	660
AGTGCACAA	CATCAGGAGG	CGCGATCTCT	GCTGAAGGGA	ACCTTGTGAT	CTCCAATAAC	720
CAAAATATCT	TTTTCGATGG	CTGCAAAGCA	ACTACAAATG	GCGGAGCTAT	TGATTGTAAC	780
AAAGCAGGGG	CGAACCCAGA	CCCTATCTTG	ACTCTTTCAG	GAAATGAGAG	CCTGCATTTT	840
CTGAATAACA	CAGCAGGAAA	TAGTGGAGGT	GCGATTTATA	CCAAAAAATT	GGTGTATCC	900
TCAGGACGAG	GAGGAGTGTT	ATTTTCTAAC	AACAAAGCTG	CGAATGCTAC	TCCTAAAGGA	960
GGGGCAATTG	CGATTCTAGA	TTCTGGAGAG	ATTAGCATTT	CTGCAGATCT	CGGCAATATC	1020
ATTTTCGAGG	GCAATACTAC	GAGCACTACA	GGAAGTCCTG	CGAGTGTGAC	CAGAAATGCT	1080
ATAGATCTTG	CATCGAATGC	AAAATTTTTA	AATCTCCGAG	CGACTCGGGG	AAATAAAGTT	1140
ATTTTCTATG	ATCCTATCAC	GAGCTCAGGA	GCTACTGATA	AGCTCTCTTT	GAATAAAGCT	1200
GACGCAGGAT	CTGGAAATAC	CTATGAAGGC	TACATCGTTT	TCTCTGGAGA	GAAACTCTCA	1260
GAAGAGGAAC	TTAAGAAACC	TGACAATCTG	AAGTCTACAT	TTACACAGGC	TGTAGAGCTT	1320
GCTGCAGGTG	CCTTAGTATT	GAAAGATGGA	GTGACTGTAG	TTGCAAATAC	TATAACGCAG	1380
GTCGAGGGAT	CGAAAGTCGT	TATGGATGGA	GGGACTACTT	TTGAGGCAAG	CGCTGAGGGG	1440
GTCACCTCTA	ATGGCCTAGC	CATTAATATA	GATTCCTTAG	ATGGGACAAA	TAAAGCTATC	1500
ATTAAGGCCA	CGGCAGCAAG	TAAGGATGTT	GCCTTATCAG	GGCCTATCAT	GCTTGTAGAT	1560
GCTCAGGGGA	ACTATTATGA	GCATCATAAT	CTCAGTCAAC	AGCAGGTCTT	TCCTTTAATA	1620
GAGCTTTCTG	CACAAGGAAC	GATGACTACT	ACAGATATCC	CCGATACCCC	AATTCTAAAT	1680
ACTACGAATC	ACTATGGGTA	TCAAGGAACT	GGAATAATTG	TTTGGGTCGA	CGATGCAACT	1740
GCAAAAACAA	AAAATGCTAC	CTTAACCTTG	ACTAAAACAG	GATACAAGCC	GAATCCAGAA	1800
CGTCAGGGAC	CTTTGGTTCC	TAATAGCCTG	TGGGGTTCTT	TTGTCGATGT	CCGCTCCATT	1860

```

CAGAGCCTCA TGGACCGGAG CACAAGTTCG TTATCTTCGT CAACAAATTT GTGGGTATCA 1920
GGAATCGCGG ACTTTTTTGCA TGAAGATCAG AAAGGAAACC AACGTAGTTA TCGTCATTCT 1980
AGCGCGGGTT ATGCATTAGG AGGAGGATTC TTCACGGCTT CTGAAAATTT CTTTAATTTT 2040
GCTTTTTGTC AGCTTTTTTG CTACGACAAG GACCATCTTG TGGCTAAGAA CCATACCCAT 2100
GTATATGCAG GGGCAATGAG TTACCGACAC CTCGGAGAGT CTAAGACCCT CGCTAAGATT 2160
TTGTCAGGAA ATTCTGACTC CCTACCTTTT GTCTTCAATG CTCGGTTTGC TTATGGCCAT 2220
ACCGACAATA ACATGACCAC AAAGTACACT GGCTATTCTC CTGTTAAGGG AAGCTGGGGA 2280
AATGATGCCT TCGGTATAGA ATGTGGAGGA GCTATCCCGG TAGTTGCTTC AGGACGTCGG 2340
TCTTGGGTGG ATACCCACAC GCCATTTCTA AACCTAGAGA TGATCTATGC ACATCAGAAT 2400
GACTTTAAGG AAAACGGCAC AGAAGGCCGT TCTTTCCAAA GTGAAGACCT CTTCAATCTA 2460
GCGGTTCCCTG TAGGGATAAA ATTTGAGAAA TTCTCCGATA AGTCTACGTA TGATCTCTCC 2520
ATAGCTTACG TTCCCGATGT GATTCGTAAT GATCCAGGCT GCACGACAAC TCTTATGGTT 2580
TCTGGGGATT CTTGGTGCAC ATGTGGTACA AGCTTGTCTA GACAAGCTCT TCTTGTACGT 2640
GCTGGAAATC ATCATGCCTT TGCTTCAAAC TTTGAAGTTT TCAGTCAGTT TGAAGTCGAG 2700
TTGCGAGGTT CTTCTCGTAG CTATGCTATC GATCTTGGAG GAAGATTCGG ATTTTAA 2757

```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 918 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met Arg Ser Ser Phe Ser Leu Leu Leu Ile Ser Ser Ser Leu Ala Phe
 1             5             10             15

Pro Leu Leu Met Ser Val Ser Ala Asp Ala Ala Asp Leu Thr Leu Gly
      20             25             30

Ser Arg Asp Ser Tyr Asn Gly Asp Thr Ser Thr Thr Glu Phe Thr Pro
      35             40             45

Lys Ala Ala Thr Ser Asp Ala Ser Gly Thr Thr Tyr Ile Leu Asp Gly
      50             55             60

Asp Val Ser Ile Ser Gln Ala Gly Lys Gln Thr Ser Leu Thr Thr Ser
      65             70             75             80

Cys Phe Ser Asn Thr Ala Gly Asn Leu Thr Phe Leu Gly Asn Gly Phe
      85             90             95

Ser Leu His Phe Asp Asn Ile Ile Ser Ser Thr Val Ala Gly Val Val
      100             105             110

Val Ser Asn Thr Ala Ala Ser Gly Ile Thr Lys Phe Ser Gly Phe Ser
      115             120             125

Thr Leu Arg Met Leu Ala Ala Pro Arg Thr Thr Gly Lys Gly Ala Ile
      130             135             140

Lys Ile Thr Asp Gly Leu Val Phe Glu Ser Ile Gly Asn Leu Asp Gln
      145             150             155             160

```

Asn Glu Asn Ala Ser Ser Glu Asn Gly Gly Ala Ile Asn Thr Lys Thr  
 165 170 175  
 Leu Ser Leu Thr Gly Ser Thr Arg Phe Val Ala Phe Leu Gly Asn Ser  
 180 185 190  
 Ser Ser Gln Gln Gly Gly Ala Ile Tyr Ala Ser Gly Asp Ser Val Ile  
 195 200 205  
 Ser Glu Asn Ala Gly Ile Leu Ser Phe Gly Asn Asn Ser Ala Thr Thr  
 210 215 220  
 Ser Gly Gly Ala Ile Ser Ala Glu Gly Asn Leu Val Ile Ser Asn Asn  
 225 230 235 240  
 Gln Asn Ile Phe Phe Asp Gly Cys Lys Ala Thr Thr Asn Gly Gly Ala  
 245 250 255  
 Ile Asp Cys Asn Lys Ala Gly Ala Asn Pro Asp Pro Ile Leu Thr Leu  
 260 265 270  
 Ser Gly Asn Glu Ser Leu His Phe Leu Asn Asn Thr Ala Gly Asn Ser  
 275 280 285  
 Gly Gly Ala Ile Tyr Thr Lys Lys Leu Val Leu Ser Ser Gly Arg Gly  
 290 295 300  
 Gly Val Leu Phe Ser Asn Asn Lys Ala Ala Asn Ala Thr Pro Lys Gly  
 305 310 315 320  
 Gly Ala Ile Ala Ile Leu Asp Ser Gly Glu Ile Ser Ile Ser Ala Asp  
 325 330 335  
 Leu Gly Asn Ile Ile Phe Glu Gly Asn Thr Thr Ser Thr Thr Gly Ser  
 340 345 350  
 Pro Ala Ser Val Thr Arg Asn Ala Ile Asp Leu Ala Ser Asn Ala Lys  
 355 360 365  
 Phe Leu Asn Leu Arg Ala Thr Arg Gly Asn Lys Val Ile Phe Tyr Asp  
 370 375 380  
 Pro Ile Thr Ser Ser Gly Ala Thr Asp Lys Leu Ser Leu Asn Lys Ala  
 385 390 395 400  
 Asp Ala Gly Ser Gly Asn Thr Tyr Glu Gly Tyr Ile Val Phe Ser Gly  
 405 410 415  
 Glu Lys Leu Ser Glu Glu Glu Leu Lys Lys Pro Asp Asn Leu Lys Ser  
 420 425 430  
 Thr Phe Thr Gln Ala Val Glu Leu Ala Ala Gly Ala Leu Val Leu Lys  
 435 440 445  
 Asp Gly Val Thr Val Val Ala Asn Thr Ile Thr Gln Val Glu Gly Ser  
 450 455 460  
 Lys Val Val Met Asp Gly Gly Thr Thr Phe Glu Ala Ser Ala Glu Gly

465		470		475		480
Val Thr Leu Asn Gly	Leu Ala Ile Asn Ile Asp Ser Leu Asp Gly Thr					
	485		490		495	
Asn Lys Ala Ile Ile Lys Ala Thr Ala Ala Ser Lys Asp Val Ala Leu						
	500		505		510	
Ser Gly Pro Ile Met Leu Val Asp Ala Gln Gly Asn Tyr Tyr Glu His						
	515		520		525	
His Asn Leu Ser Gln Gln Gln Val Phe Pro Leu Ile Glu Leu Ser Ala						
	530		535		540	
Gln Gly Thr Met Thr Thr Thr Asp Ile Pro Asp Thr Pro Ile Leu Asn						
	545		550		555	560
Thr Thr Asn His Tyr Gly Tyr Gln Gly Thr Gly Ile Ile Val Trp Val						
	565		570			575
Asp Asp Ala Thr Ala Lys Thr Lys Asn Ala Thr Leu Thr Trp Thr Lys						
	580		585		590	
Thr Gly Tyr Lys Pro Asn Pro Glu Arg Gln Gly Pro Leu Val Pro Asn						
	595		600		605	
Ser Leu Trp Gly Ser Phe Val Asp Val Arg Ser Ile Gln Ser Leu Met						
	610		615		620	
Asp Arg Ser Thr Ser Ser Leu Ser Ser Ser Thr Asn Leu Trp Val Ser						
	625		630		635	640
Gly Ile Ala Asp Phe Leu His Glu Asp Gln Lys Gly Asn Gln Arg Ser						
	645		650		655	
Tyr Arg His Ser Ser Ala Gly Tyr Ala Leu Gly Gly Gly Phe Phe Thr						
	660		665		670	
Ala Ser Glu Asn Phe Phe Asn Phe Ala Phe Cys Gln Leu Phe Gly Tyr						
	675		680		685	
Asp Lys Asp His Leu Val Ala Lys Asn His Thr His Val Tyr Ala Gly						
	690		695		700	
Ala Met Ser Tyr Arg His Leu Gly Glu Ser Lys Thr Leu Ala Lys Ile						
	705		710		715	720
Leu Ser Gly Asn Ser Asp Ser Leu Pro Phe Val Phe Asn Ala Arg Phe						
	725		730		735	
Ala Tyr Gly His Thr Asp Asn Asn Met Thr Thr Lys Tyr Thr Gly Tyr						
	740		745		750	
Ser Pro Val Lys Gly Ser Trp Gly Asn Asp Ala Phe Gly Ile Glu Cys						
	755		760		765	
Gly Gly Ala Ile Pro Val Val Ala Ser Gly Arg Arg Ser Trp Val Asp						
	770		775		780	

Thr His Thr Pro Phe Leu Asn Leu Glu Met Ile Tyr Ala His Gln Asn  
 785 790 795 800  
 Asp Phe Lys Glu Asn Gly Thr Glu Gly Arg Ser Phe Gln Ser Glu Asp  
 805 810 815  
 Leu Phe Asn Leu Ala Val Pro Val Gly Ile Lys Phe Glu Lys Phe Ser  
 820 825 830  
 Asp Lys Ser Thr Tyr Asp Leu Ser Ile Ala Tyr Val Pro Asp Val Ile  
 835 840 845  
 Arg Asn Asp Pro Gly Cys Thr Thr Thr Leu Met Val Ser Gly Asp Ser  
 850 855 860  
 Trp Ser Thr Cys Gly Thr Ser Leu Ser Arg Gln Ala Leu Leu Val Arg  
 865 870 875 880  
 Ala Gly Asn His His Ala Phe Ala Ser Asn Phe Glu Val Phe Ser Gln  
 885 890 895  
 Phe Glu Val Glu Leu Arg Gly Ser Ser Arg Ser Tyr Ala Ile Asp Leu  
 900 905 910  
 Gly Gly Arg Phe Gly Phe  
 915

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2787 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATGAAATCCT	CTCTTCATTG	GTTTGTAATC	TCGTCATCTT	TAGCACTTCC	CTTGTCACTA	60
AATTTCTCTG	CGTTTGCTGC	TGTTGTTGAA	ATCAATCTAG	GACCTACCAA	TAGCTTCTCT	120
GGACCAGGAA	CCTACACTCC	TCCAGCCCAA	ACAACAAATG	CAGATGGAAC	TATCTATAAT	180
CTAACAGGGG	ATGTCTCAAT	CACCAATGCA	GGATCTCCGA	CAGCTCTAAC	CGCTTCCTGC	240
TTTAAAGAAA	CTACTGGGAA	TCTTTCTTTC	CAAGGCCACG	GCTACCAATT	TCTCCTACAA	300
AATATCGATG	CGGGAGCGAA	CTGTACCTTT	ACCAATACAG	CTGCAAATAA	GCTTCTCTCC	360
TTTTTCAGGAT	TCTCCTATTT	GTCACTAATA	CAAACCACGA	ATGCTACCAC	AGGAACAGGA	420
GCCATCAAGT	CCACAGGAGC	TTGTTCTATT	CAGTCGAACT	ATAGTTGCTA	CTTTGGCCAA	480
AACTTTTCTA	ATGACAATGG	AGGCGCCCTC	CAAGGCAGCT	CTATCAGTCT	ATCGCTAAAC	540
CCCAACCTAA	CGTTTGCCAA	AAACAAAGCA	ACGCAAAAAG	GGGGTGCCCT	CTATTCCACG	600
GGAGGGATTA	CAATTAACAA	TACGTAAAC	TCAGCATCAT	TTTCTGAAAA	TACCGCGGCG	660
AACAATGGCG	GAGCCATTTA	CACGGAAGCT	AGCAGTTTTA	TTAGCAGCAA	CAAAGCAATT	720
AGCTTTATAA	ACAATAGTGT	GACCGCAACC	TCAGCTACAG	GGGGAGCCAT	TTACTGTAGT	780
AGTACATCAG	CCCCCAAACC	AGTCTTAACT	CTATCAGACA	ACGGGGAAC	GAAGTTTATA	840
GGAAATACAG	CAATTACTAG	TGGTGGGGCG	ATTTATACTG	ACAATCTAGT	TCTTTCTTCT	900
GGAGGACCTA	CGCTTTTTTA	AAACAACTCT	GCTATAGATA	CTGCAGCTCC	CTTAGGAGGA	960
GCAATTGCGA	TTGCTGACTC	TGGATCTTTG	AGTCTTTCGG	CTCTTGGTGG	AGACATCACT	1020
TTTGAAGGAA	ACACAGTAGT	CAAAGGAGCT	TCTTCGAGTC	AGACCACTAC	CAGAAATTCT	1080

ATTAACATCG	GAAACACCAA	TGCTAAGATT	GTACAGCTGC	GAGCCTCTCA	AGGCAATACT	1140
ATCTACTTCT	ATGATCCTAT	AACAACCTAAC	CATACTGCAG	CTCTCTCAGA	TGCTCTAAAC	1200
TTAAATGGTC	CTGACCTTGC	AGGGAATCCT	GCATATCAAG	GAACCATCGT	ATTTTCTGGA	1260
GAGAAGCTCT	CGGAAGCAGA	AGCTGCAGAA	GCTGATAATC	TCAAATCTAC	AATTCAGCAA	1320
CCTCTAACTC	TTGCGGGAGG	GCAACTCTCT	CTTAAATCAG	GAGTCACTCT	AGTTGCTAAG	1380
TCCTTTTTCG	AATCTCCGGG	CTCTACCCCT	CTCATGGATG	CAGGGACCAC	ATTAGAAACC	1440
GCTGATGGGA	TCACTATCAA	TAATCTTGTT	CTCAATGTAG	ATTCTTTAAA	AGAGACCAAG	1500
AAGGCTACGC	TAAAGACAAC	ACAAGCAAGT	CAGACAGTCA	CTTTATCTGG	ATCGCTCTCT	1560
CTTGCTAGATC	CTTCTGGAAA	TGTCTACGAA	GATGTCTCTT	GGAATAACCC	TCAAGTCTTT	1620
TCTTGTCTCA	CTCTTACTGC	TGACGACCCC	GCGAATATTC	ACATCACAGA	CTTAGCTGCT	1680
GATCCCCTAG	AAAAAATCC	TATCCATTGG	GGATACCAAG	GGAATTGGGC	ATTATCTTGG	1740
CAAGAGGATA	CTGCGACTAA	ATCCAAAGCA	GCGACTCTTA	CCTGGACAAA	AACAGGATAC	1800
AATCCGAATC	CTGAGCGTCG	TGGAACCTTA	GTTGCTAACA	CGCTATGGGG	ATCCTTTGTT	1860
GATGTGCGCT	CCATACAACA	GCTTGTAGCC	ACTAAAGTAC	GCCAATCTCA	AGAAACTCGC	1920
GGCATCTGGT	GTGAAGGGAT	CTCGAACTTC	TTCCATAAAG	ATAGCACGAA	GATAAATAAA	1980
GGTTTTTCGCC	ACATAAGTGC	AGGTTATGTT	GTAGGAGCGA	CTACAACATT	AGCTTCTGAT	2040
AATCTTATCA	CTGCAGCCTT	CTGCCAATTA	TTCCGGGAAAG	ATAGAGATCA	CTTTATAAAT	2100
AAAAATAGAG	CTTCTGCCTA	TGCAGCTTCT	CTCCATCTCC	AGCATCTAGC	GACCTTGTCT	2160
TCTCCAAGCT	TGTTACGCTA	CCTTCCTGGA	TCTGAAAGTG	AGCAGCCTGT	CCTCTTTGAT	2220
GCTCAGATCA	GCTATATCTA	TAGTAAAAAT	ACTATGAAAA	CCTATTACAC	CCAAGCACCA	2280
AAGGGAGAGA	GCTCGTGGTA	TAATGACGGT	TGCGCTCTGG	AACTTGCGAG	CTCCCTACCA	2340
CACACTGCTT	TAAGCCATGA	GGGTCTCTTC	CACGCGTATT	TTCCCTTTCAT	CAAAGTAGAA	2400
GCTTCGTACA	TACACCAAGA	TAGCTTCAAA	GAACGTAATA	CTACCTTGGT	ACGATCTTTC	2460
GATAGCGGTG	ATTTAATTAA	CGTCTCTGTG	CCTATTGGAA	TTACCTTCGA	GAGATTCTCG	2520
AGAAACGAGC	GTGCGTCTTA	CGAAGCTACT	GTCATCTACG	TTGCCGATGT	CTATCGTAAG	2580
AATCCTGACT	GCACGACAGC	TCTCCTAATC	AACAATACCT	CGTGGAACAC	TACAGCAACG	2640
AATCTCTCAA	GACAAGCTGG	TATCGGAAGA	GCAGGGATCT	TTTATGCCTT	CTCTCCAAAT	2700
CTTGAGGTCA	CAAGTAACCT	ATCTATGGAA	ATTCGTGGAT	CTTCACGCAG	CTACAATGCA	2760
GATCTTGGAG	GTAAGTTCCA	GTTCTAA				2787

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 928 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Lys	Ser	Ser	Leu	His	Trp	Phe	Val	Ile	Ser	Ser	Ser	Leu	Ala	Leu
1				5					10					15	
Pro	Leu	Ser	Leu	Asn	Phe	Ser	Ala	Phe	Ala	Ala	Val	Val	Glu	Ile	Asn
			20					25					30		
Leu	Gly	Pro	Thr	Asn	Ser	Phe	Ser	Gly	Pro	Gly	Thr	Tyr	Thr	Pro	Pro
		35					40					45			
Ala	Gln	Thr	Thr	Asn	Ala	Asp	Gly	Thr	Ile	Tyr	Asn	Leu	Thr	Gly	Asp
		50				55					60				
Val	Ser	Ile	Thr	Asn	Ala	Gly	Ser	Pro	Thr	Ala	Leu	Thr	Ala	Ser	Cys
	65				70				75						80
Phe	Lys	Glu	Thr	Thr	Gly	Asn	Leu	Ser	Phe	Gln	Gly	His	Gly	Tyr	Gln

85										90					95				
Phe	Leu	Leu	Gln	Asn	Ile	Asp	Ala	Gly	Ala	Asn	Cys	Thr	Phe	Thr	Asn				
			100					105					110						
Thr	Ala	Ala	Asn	Lys	Leu	Leu	Ser	Phe	Ser	Gly	Phe	Ser	Tyr	Leu	Ser				
		115					120					125							
Leu	Ile	Gln	Thr	Thr	Asn	Ala	Thr	Thr	Gly	Thr	Gly	Ala	Ile	Lys	Ser				
	130					135					140								
Thr	Gly	Ala	Cys	Ser	Ile	Gln	Ser	Asn	Tyr	Ser	Cys	Tyr	Phe	Gly	Gln				
145					150					155					160				
Asn	Phe	Ser	Asn	Asp	Asn	Gly	Gly	Ala	Leu	Gln	Gly	Ser	Ser	Ile	Ser				
				165					170					175					
Leu	Ser	Leu	Asn	Pro	Asn	Leu	Thr	Phe	Ala	Lys	Asn	Lys	Ala	Thr	Gln				
			180					185					190						
Lys	Gly	Gly	Ala	Leu	Tyr	Ser	Thr	Gly	Gly	Ile	Thr	Ile	Asn	Asn	Thr				
	195						200					205							
Leu	Asn	Ser	Ala	Ser	Phe	Ser	Glu	Asn	Thr	Ala	Ala	Asn	Asn	Gly	Gly				
	210					215					220								
Ala	Ile	Tyr	Thr	Glu	Ala	Ser	Ser	Phe	Ile	Ser	Ser	Asn	Lys	Ala	Ile				
225					230					235					240				
Ser	Phe	Ile	Asn	Asn	Ser	Val	Thr	Ala	Thr	Ser	Ala	Thr	Gly	Gly	Ala				
			245					250					255						
Ile	Tyr	Cys	Ser	Ser	Thr	Ser	Ala	Pro	Lys	Pro	Val	Leu	Thr	Leu	Ser				
			260					265					270						
Asp	Asn	Gly	Glu	Leu	Asn	Phe	Ile	Gly	Asn	Thr	Ala	Ile	Thr	Ser	Gly				
		275					280					285							
Gly	Ala	Ile	Tyr	Thr	Asp	Asn	Leu	Val	Leu	Ser	Ser	Gly	Gly	Pro	Thr				
	290					295					300								
Leu	Phe	Lys	Asn	Asn	Ser	Ala	Ile	Asp	Thr	Ala	Ala	Pro	Leu	Gly	Gly				
305					310					315				320					
Ala	Ile	Ala	Ile	Ala	Asp	Ser	Gly	Ser	Leu	Ser	Leu	Ser	Ala	Leu	Gly				
				325					330					335					
Gly	Asp	Ile	Thr	Phe	Glu	Gly	Asn	Thr	Val	Val	Lys	Gly	Ala	Ser	Ser				
			340					345					350						
Ser	Gln	Thr	Thr	Thr	Arg	Asn	Ser	Ile	Asn	Ile	Gly	Asn	Thr	Asn	Ala				
			355				360					365							
Lys	Ile	Val	Gln	Leu	Arg	Ala	Ser	Gln	Gly	Asn	Thr	Ile	Tyr	Phe	Tyr				
	370					375					380								
Asp	Pro	Ile	Thr	Thr	Asn	His	Thr	Ala	Ala	Leu	Ser	Asp	Ala	Leu	Asn				
385					390					395					400				



Leu Asn Gly Pro Asp Leu Ala Gly Asn Pro Ala Tyr Gln Gly Thr Ile  
 405 410 415  
 Val Phe Ser Gly Glu Lys Leu Ser Glu Ala Glu Ala Ala Glu Ala Asp  
 420 425 430  
 Asn Leu Lys Ser Thr Ile Gln Gln Pro Leu Thr Leu Ala Gly Gly Gln  
 435 440 445  
 Leu Ser Leu Lys Ser Gly Val Thr Leu Val Ala Lys Ser Phe Ser Gln  
 450 455 460  
 Ser Pro Gly Ser Thr Leu Leu Met Asp Ala Gly Thr Thr Leu Glu Thr  
 465 470 475 480  
 Ala Asp Gly Ile Thr Ile Asn Asn Leu Val Leu Asn Val Asp Ser Leu  
 485 490 495  
 Lys Glu Thr Lys Lys Ala Thr Leu Lys Ala Thr Gln Ala Ser Gln Thr  
 500 505 510  
 Val Thr Leu Ser Gly Ser Leu Ser Leu Val Asp Pro Ser Gly Asn Val  
 515 520 525  
 Tyr Glu Asp Val Ser Trp Asn Asn Pro Gln Val Phe Ser Cys Leu Thr  
 530 535 540  
 Leu Thr Ala Asp Asp Pro Ala Asn Ile His Ile Thr Asp Leu Ala Ala  
 545 550 555 560  
 Asp Pro Leu Glu Lys Asn Pro Ile His Trp Gly Tyr Gln Gly Asn Trp  
 565 570 575  
 Ala Leu Ser Trp Gln Glu Asp Thr Ala Thr Lys Ser Lys Ala Ala Thr  
 580 585 590  
 Leu Thr Trp Thr Lys Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg Gly  
 595 600 605  
 Thr Leu Val Ala Asn Thr Leu Trp Gly Ser Phe Val Asp Val Arg Ser  
 610 615 620  
 Ile Gln Gln Leu Val Ala Thr Lys Val Arg Gln Ser Gln Glu Thr Arg  
 625 630 635 640  
 Gly Ile Trp Cys Glu Gly Ile Ser Asn Phe Phe His Lys Asp Ser Thr  
 645 650 655  
 Lys Ile Asn Lys Gly Phe Arg His Ile Ser Ala Gly Tyr Val Val Gly  
 660 665 670  
 Ala Thr Thr Thr Leu Ala Ser Asp Asn Leu Ile Thr Ala Ala Phe Cys  
 675 680 685  
 Gln Leu Phe Gly Lys Asp Arg Asp His Phe Ile Asn Lys Asn Arg Ala  
 690 695 700

Ser Ala Tyr Ala Ala Ser Leu His Leu Gln His Leu Ala Thr Leu Ser  
 705 710 715 720  
 Ser Pro Ser Leu Leu Arg Tyr Leu Pro Gly Ser Glu Ser Glu Gln Pro  
 725 730 735  
 Val Leu Phe Asp Ala Gln Ile Ser Tyr Ile Tyr Ser Lys Asn Thr Met  
 740 745 750  
 Lys Thr Tyr Tyr Thr Gln Ala Pro Lys Gly Glu Ser Ser Trp Tyr Asn  
 755 760 765  
 Asp Gly Cys Ala Leu Glu Leu Ala Ser Ser Leu Pro His Thr Ala Leu  
 770 775 780  
 Ser His Glu Gly Leu Phe His Ala Tyr Phe Pro Phe Ile Lys Val Glu  
 785 790 795 800  
 Ala Ser Tyr Ile His Gln Asp Ser Phe Lys Glu Arg Asn Thr Thr Leu  
 805 810 815  
 Val Arg Ser Phe Asp Ser Gly Asp Leu Ile Asn Val Ser Val Pro Ile  
 820 825 830  
 Gly Ile Thr Phe Glu Arg Phe Ser Arg Asn Glu Arg Ala Ser Tyr Glu  
 835 840 845  
 Ala Thr Val Ile Tyr Val Ala Asp Val Tyr Arg Lys Asn Pro Asp Cys  
 850 855 860  
 Thr Thr Ala Leu Leu Ile Asn Asn Thr Ser Trp Lys Thr Thr Gly Thr  
 865 870 875 880  
 Asn Leu Ser Arg Gln Ala Gly Ile Gly Arg Ala Gly Ile Phe Tyr Ala  
 885 890 895  
 Phe Ser Pro Asn Leu Glu Val Thr Ser Asn Leu Ser Met Glu Ile Arg  
 900 905 910  
 Gly Ser Ser Arg Ser Tyr Asn Ala Asp Leu Gly Gly Lys Phe Gln Phe  
 915 920 925

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2793 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATGAAAATAC CTTGCACAA ACTCCTGATC TCTTCGACTC TTGTCACTCC CATTCTATTG	60
AGCATTTGCAA CTTACGGAGC AGATGCTTCT TTATCCCCTA CAGATAGCTT TGATGGAGCG	120
GGCGGCTCTA CATTCTACTCC AAAATCTACA GCAGATGCCA ATGGAACGAA CTATGTCTTA	180
TCAGGAAATG TCTATATAAA CGATGCTGGG AAAGGCACAG CATTAAACAGG CTGCTGCTTT	240
ACAGAAACTA CGGGTGATCT GACATTTACT GGAAAGGGAT ACTCATTTTC ATTCAACACG	300

GTAGATGCGG	GTTCGAATGC	AGGAGCTGCG	GCAAGCACAA	CTGCTGATAA	AGCCCTAACA	360
TTCACAGGAT	TTTCTAACCT	TTCCTTCATT	GCAGCTCCTG	GAACCTACAGT	TGCTTCAGGA	420
AAAAGTACTT	TAAGTTCTGC	AGGAGCCTTA	AATCTTACCG	ATAATGGAAC	GATTCTCTTT	480
AGCCAAAACG	TCTCCAATGA	AGCTAATAAC	AATGGCGGAG	CGATCACCAC	AAAAACTCTT	540
TCTATTTCTG	GGAATACCTC	TTCTATAACC	TTCAC TAGTA	ATAGCGCAA	AAAATTAGGT	600
GGAGCGATCT	ATAGCTCTGC	GGCTGCAAGT	ATTT CAGGAA	ACACCGGCCA	GTTAGTCTTT	660
ATGAATAATA	AAGGAGAAAC	TGGGGGCGGG	GCTCTGGGCT	TTGAAGCCAG	CTCCTCGATT	720
ACTCAAAATA	GTCCTCTTTT	CTTCTCTGGA	AACACTGCAA	CAGATGCTGC	AGGCAAGGGC	780
GGGGCCATTT	ATTGTGAAAA	AACAGGAGAG	ACTCCTACTC	TTACTATCTC	TGGAAATAAA	840
AGTCTGACCT	TCGCCGAGAA	CTCTTCAGTA	ACTCAAGGCG	GAGCAATCTG	TGCCCATGGT	900
CTAGATCTTT	CCGCTGCTGG	CCCTACCCTA	TTTTCAAATA	ATAGATGCGG	GAACACAGCT	960
GCAGGCAAGG	GCGGCGCTAT	TGCAATTGCC	GACTCTGGAT	CTTTAAGTCT	CTCTGCAAAT	1020
CAAGGAGACA	TCACGTTCTT	TGGCAACACT	CTAACCTCAA	CCTCCGCGCC	AACATCGACA	1080
CGGAATGCTA	TCTACCTGGG	ATCGTCAGCA	AAAATTACGA	ACTTAAGGGC	AGCCCAAGGC	1140
CAATCTATCT	ATTTCTATGA	TCCGATTGCA	TCTAACACCA	CAGGAGCTTC	AGACGTTCTG	1200
ACCATCAACC	AACCGGATAG	CAACTCGCCT	TTAGATTATT	CAGGAACGAT	TGTATTTTCT	1260
GGGGAAAAGC	TCTCTGCAGA	TGAAGCGAAA	GCTGCTGATA	ACTTCACATC	TATATTAAAG	1320
CAACCATTTG	CTCTAGCCTC	TGGAACCTTA	GCACTCAAAG	GAAATGTCTG	GTTAGATGTC	1380
AATGGTTTCA	CACAGACTGA	AGGCTCTACA	CTCCTCATGC	AACCAGGAAC	AAAGCTCAAA	1440
GCAGATACTG	AAGCTATCAG	TCTTACCAAA	CTTGTCGTTG	ATCTTTCTGC	CTTAGAGGGA	1500
AATAAGAGTG	TGTCCATTGA	AACAGCAGGA	GCCAACAAAA	CTATAACTCT	AACCTCTCCT	1560
CTTGTTTTCC	AAGATAGTAG	CGGCAATTTT	TATGAAAGCC	ATACGATAAA	CCAAGCCTTC	1620
ACGCAGCCTT	TGGTGGTATT	CACTGCTGCT	ACTGCTGCTA	GCGATATTTA	TATCGATGCG	1680
CTTCTCACTT	CTCCAGTACA	AACTCCAGAA	CCTCATTACG	GGTATCAGGG	ACATTGGGAA	1740
GCCACTTGCG	CAGACACATC	AACTGCAAAA	TCAGGAACTA	TGACTTGGGT	AACTACGGGC	1800
TACAACCCTA	ATCCTGAGCG	TAGAGCTTCC	GTAGTTCCCG	ATTCATTATG	GGCATCCTTT	1860
ACTGACATTC	GCACTCTACA	GCAGATCATG	ACATCTCAAG	CGAATAGTAT	CTATCAGCAA	1920
CGAGGACTCT	GGGCATCAGG	AACTGCGAAT	TTCTTCCATA	AGGATAAATC	AGGAAC TAAC	1980
CAAGCATTCC	GACATAAAAG	CTACGGCTAT	ATTGTTGGAG	GAAGTGCTGA	AGATTTTCT	2040
GAAAATATCT	TCAGTGTAGC	TTTCTGCCAG	CTCTTCGGTA	AAGATAAAGA	CCTGTTTATA	2100
GTTGAAAATA	CCTCTCATAA	CTATTTAGCG	TCGCTATACC	TGCAACATCG	AGCATTCCTA	2160
GGAGGACTTC	CCATGCCCTC	ATTTGGAAGT	ATCACCGACA	TGCTGAAAGA	TATTCTCTC	2220
ATTTTGAAATG	CCCAGCTAAG	CTACAGCTAC	ACTAAAAATG	ATATGGATAC	TCGCTATACT	2280
TCCTATCCTG	AAGCTCAAGG	TTCTTGGACC	AATAATTCTG	GGGCTCTAGA	GCTCGGAGGA	2340
TCTCTGGCTC	TATATCTCCC	TAAAGAAGCA	CCGTTCTTCC	AGGGATATTT	CCCCTTCTTA	2400
AAGTTCCAGG	CAGTCTACAG	CCGCCAACAA	AACTTTAAAG	AGAGTGGCGC	TGAAGCCCGT	2460
GCTTTTGATG	ATGGAGACCT	AGTGAAGTGC	TCTATCCCTG	TCGGCATTCG	GTTAGAAAAA	2520
ATCTCCGAAG	ATGAAAAAAA	TAATTTTCGAG	ATTTCTCTAG	CCAACATTGG	TGATGTGTAT	2580
CGTAAAAATC	CCGTTTCGCG	TACTTCTCTA	ATGGTCAGTG	GAGCCTCTTG	GACTTCGCTA	2640
TGTAAAAACC	TCGCACGACA	AGCCTTCTTA	GCAAGTGCTG	GAAGCCATCT	GACTCTCTCC	2700
CCTCATGTAG	AACTCTCTGG	GGAAGCTGCT	TATGAGCTTC	GTGGCTCAGC	ACACATCTAC	2760
AATGTAGATT	GTGGGCTAAG	ATACTCATTC	TAG			2793

## (2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 930 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met	Lys	Ile	Pro	Leu	His	Lys	Leu	Leu	Ile	Ser	Ser	Thr	Leu	Val	Thr
1							5								15
									10						

Pro Ile Leu Leu Ser Ile Ala Thr Tyr Gly Ala Asp Ala Ser Leu Ser  
                   20                  25                  30

Pro Thr Asp Ser Phe Asp Gly Ala Gly Gly Ser Thr Phe Thr Pro Lys  
                   35                  40                  45

Ser Thr Ala Asp Ala Asn Gly Thr Asn Tyr Val Leu Ser Gly Asn Val  
           50                  55                  60

Tyr Ile Asn Asp Ala Gly Lys Gly Thr Ala Leu Thr Gly Cys Cys Phe  
   65                  70                  75                  80

Thr Glu Thr Thr Gly Asp Leu Thr Phe Thr Gly Lys Gly Tyr Ser Phe  
                   85                  90                  95

Ser Phe Asn Thr Val Asp Ala Gly Ser Asn Ala Gly Ala Ala Ala Ser  
                   100                  105                  110

Thr Thr Ala Asp Lys Ala Leu Thr Phe Thr Gly Phe Ser Asn Leu Ser  
           115                  120                  125

Phe Ile Ala Ala Pro Gly Thr Thr Val Ala Ser Gly Lys Ser Thr Leu  
   130                  135                  140

Ser Ser Ala Gly Ala Leu Asn Leu Thr Asp Asn Gly Thr Ile Leu Phe  
  145                  150                  155                  160

Ser Gln Asn Val Ser Asn Glu Ala Asn Asn Asn Gly Gly Ala Ile Thr  
                   165                  170                  175

Thr Lys Thr Leu Ser Ile Ser Gly Asn Thr Ser Ser Ile Thr Phe Thr  
           180                  185                  190

Ser Asn Ser Ala Lys Lys Leu Gly Gly Ala Ile Tyr Ser Ser Ala Ala  
           195                  200                  205

Ala Ser Ile Ser Gly Asn Thr Gly Gln Leu Val Phe Met Asn Asn Lys  
   210                  215                  220

Gly Glu Thr Gly Gly Gly Ala Leu Gly Phe Glu Ala Ser Ser Ser Ile  
  225                  230                  235                  240

Thr Gln Asn Ser Ser Leu Phe Phe Ser Gly Asn Thr Ala Thr Asp Ala  
                   245                  250                  255

Ala Gly Lys Gly Gly Ala Ile Tyr Cys Glu Lys Thr Gly Glu Thr Pro  
           260                  265                  270

Thr Leu Thr Ile Ser Gly Asn Lys Ser Leu Thr Phe Ala Glu Asn Ser  
           275                  280                  285

Ser Val Thr Gln Gly Gly Ala Ile Cys Ala His Gly Leu Asp Leu Ser  
   290                  295                  300

Ala Ala Gly Pro Thr Leu Phe Ser Asn Asn Arg Cys Gly Asn Thr Ala  
  305                  310                  315                  320

Ala Gly Lys Gly Gly Ala Ile Ala Ile Ala Asp Ser Gly Ser Leu Ser  
           325                  330                  335

Leu Ser Ala Asn Gln Gly Asp Ile Thr Phe Leu Gly Asn Thr Leu Thr  
 340 345 350  
 Ser Thr Ser Ala Pro Thr Ser Thr Arg Asn Ala Ile Tyr Leu Gly Ser  
 355 360 365  
 Ser Ala Lys Ile Thr Asn Leu Arg Ala Ala Gln Gly Gln Ser Ile Tyr  
 370 375 380  
 Phe Tyr Asp Pro Ile Ala Ser Asn Thr Thr Gly Ala Ser Asp Val Leu  
 385 390 395 400  
 Thr Ile Asn Gln Pro Asp Ser Asn Ser Pro Leu Asp Tyr Ser Gly Thr  
 405 410 415  
 Ile Val Phe Ser Gly Glu Lys Leu Ser Ala Asp Glu Ala Lys Ala Ala  
 420 425 430  
 Asp Asn Phe Thr Ser Ile Leu Lys Gln Pro Leu Ala Leu Ala Ser Gly  
 435 440 445  
 Thr Leu Ala Leu Lys Gly Asn Val Glu Leu Asp Val Asn Gly Phe Thr  
 450 455 460  
 Gln Thr Glu Gly Ser Thr Leu Leu Met Gln Pro Gly Thr Lys Leu Lys  
 465 470 475 480  
 Ala Asp Thr Glu Ala Ile Ser Leu Thr Lys Leu Val Val Asp Leu Ser  
 485 490 495  
 Ala Leu Glu Gly Asn Lys Ser Val Ser Ile Glu Thr Ala Gly Ala Asn  
 500 505 510  
 Lys Thr Ile Thr Leu Thr Ser Pro Leu Val Phe Gln Asp Ser Ser Gly  
 515 520 525  
 Asn Phe Tyr Glu Ser His Thr Ile Asn Gln Ala Phe Thr Gln Pro Leu  
 530 535 540  
 Val Val Phe Thr Ala Ala Thr Ala Ala Ser Asp Ile Tyr Ile Asp Ala  
 545 550 555 560  
 Leu Leu Thr Ser Pro Val Gln Thr Pro Glu Pro His Tyr Gly Tyr Gln  
 565 570 575  
 Gly His Trp Glu Ala Thr Trp Ala Asp Thr Ser Thr Ala Lys Ser Gly  
 580 585 590  
 Thr Met Thr Trp Val Thr Thr Gly Tyr Asn Pro Asn Pro Glu Arg Arg  
 595 600 605  
 Ala Ser Val Val Pro Asp Ser Leu Trp Ala Ser Phe Thr Asp Ile Arg  
 610 615 620  
 Thr Leu Gln Gln Ile Met Thr Ser Gln Ala Asn Ser Ile Tyr Gln Gln  
 625 630 635 640

Arg Gly Leu Trp Ala Ser Gly Thr Ala Asn Phe Phe His Lys Asp Lys  
 645 650 655  
 Ser Gly Thr Asn Gln Ala Phe Arg His Lys Ser Tyr Gly Tyr Ile Val  
 660 665 670  
 Gly Gly Ser Ala Glu Asp Phe Ser Glu Asn Ile Phe Ser Val Ala Phe  
 675 680 685  
 Cys Gln Leu Phe Gly Lys Asp Lys Asp Leu Phe Ile Val Glu Asn Thr  
 690 695 700  
 Ser His Asn Tyr Leu Ala Ser Leu Tyr Leu Gln His Arg Ala Phe Leu  
 705 710 715 720  
 Gly Gly Leu Pro Met Pro Ser Phe Gly Ser Ile Thr Asp Met Leu Lys  
 725 730 735  
 Asp Ile Pro Leu Ile Leu Asn Ala Gln Leu Ser Tyr Ser Tyr Thr Lys  
 740 745 750  
 Asn Asp Met Asp Thr Arg Tyr Thr Ser Tyr Pro Glu Ala Gln Gly Ser  
 755 760 765  
 Trp Thr Asn Asn Ser Gly Ala Leu Glu Leu Gly Gly Ser Leu Ala Leu  
 770 775 780  
 Tyr Leu Pro Lys Glu Ala Pro Phe Phe Gln Gly Tyr Phe Pro Phe Leu  
 785 790 795 800  
 Lys Phe Gln Ala Val Tyr Ser Arg Gln Gln Asn Phe Lys Glu Ser Gly  
 805 810 815  
 Ala Glu Ala Arg Ala Phe Asp Asp Gly Asp Leu Val Asn Cys Ser Ile  
 820 825 830  
 Pro Val Gly Ile Arg Leu Glu Lys Ile Ser Glu Asp Glu Lys Asn Asn  
 835 840 845  
 Phe Glu Ile Ser Leu Ala Asn Ile Gly Asp Val Tyr Arg Lys Asn Pro  
 850 855 860  
 Arg Ser Arg Thr Ser Leu Met Val Ser Gly Ala Ser Trp Thr Ser Leu  
 865 870 875 880  
 Cys Lys Asn Leu Ala Arg Gln Ala Phe Leu Ala Ser Ala Gly Ser His  
 885 890 895  
 Leu Thr Leu Ser Pro His Val Glu Leu Ser Gly Glu Ala Ala Tyr Glu  
 900 905 910  
 Leu Arg Gly Ser Ala His Ile Tyr Asn Val Asp Cys Gly Leu Arg Tyr  
 915 920 925  
 Ser Phe  
 930

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 840 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

```
GAAGACAATA TAAGGTACCG TCATAACAGC GGGGGTTATG CACTAGGGAT CACAGCAACA      60
ACTCCTGCCG AGGATCAGCT TACTTTTGCC TTCTGCCAGC TCTTTGCTAG AGATCGCAAT     120
CATATTACAG GTAAGAACCA CGGAGATACT TACGGTGCCT CTTTGTATTT CCACCATACA     180
GAAGGGCTCT TCGACATCGC CAATTTCTCT TGGGGAAAAG CAACCCGAGC TCCCTGGGTG     240
CTCTCTGAGA TCTCCCAGAT CATTCCCTTA TCGTTCGATG CTAAATTCAG TTATCTCCAT     300
ACAGACAACC ACATGAAGAC ATATTATACC GATAACTCTA TCATCAAGGG TTCTTGGAGA     360
AACGATGCCT TCTGTGCAGA TCTTGGAGCT AGCCTGCCTT TTGTTATTTT CGTTCCGTAT     420
CTTCTGAAAG AAGTCGAACC TTTTGTCAAA GTACAGTATA TCTATGCGCA TCAGCAAGAC     480
TTCTACGAGC GTCATGCTGA AGGACGCGCT TTCAATAAAA GCGAGCTTAT CAACGTAGAG     540
ATTCCTATAG GCGTCACCTT CGAAAGAGAC TCAAAATCAG AAAAGGGAAC TTACGATCTT     600
ACTCTTATGT ATATACTCGA TGCTTACCGA CGCAATCCTA AATGTCAAAC TTCCCTAATA     660
GCTAGCGATG CTAAGTGGAT GGCCTATGGT ACCAACCTCG CACGACAAGG TTTTCTGTGTT     720
CGTGTGCGA ACCATTTCCA AGTGAACCCC CACATGGAAA TCTTCGGTCA ATTCGCTTTT     780
GAAGTACGAA GTTCTTCACG AAATTATAAT ACAAACCTAG GCTCTAAGTT TTGTTTCTAG     840
```

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 279 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Glu Asp Asn Ile Arg Tyr Arg His Asn Ser Gly Gly Tyr Ala Leu Gly
 1              5              10              15

Ile Thr Ala Thr Thr Pro Ala Glu Asp Gln Leu Thr Phe Ala Phe Cys
      20              25              30

Gln Leu Phe Ala Arg Asp Arg Asn His Ile Thr Gly Lys Asn His Gly
      35              40              45

Asp Thr Tyr Gly Ala Ser Leu Tyr Phe His His Thr Glu Gly Leu Phe
      50              55              60

Asp Ile Ala Asn Phe Leu Trp Gly Lys Ala Thr Arg Ala Pro Trp Val
      65              70              75              80

Leu Ser Glu Ile Ser Gln Ile Ile Pro Leu Ser Phe Asp Ala Lys Phe
      85              90              95

Ser Tyr Leu His Thr Asp Asn His Met Lys Thr Tyr Tyr Thr Asp Asn
      100              105              110
```

Ser Ile Ile Lys Gly Ser Trp Arg Asn Asp Ala Phe Cys Ala Asp Leu  
 115 120 125  
 Gly Ala Ser Leu Pro Phe Val Ile Ser Val Pro Tyr Leu Leu Lys Glu  
 130 135 140  
 Val Glu Pro Phe Val Lys Val Gln Tyr Ile Tyr Ala His Gln Gln Asp  
 145 150 155 160  
 Phe Tyr Glu Arg His Ala Glu Gly Arg Ala Phe Asn Lys Ser Glu Leu  
 165 170 175  
 Ile Asn Val Glu Ile Pro Ile Gly Val Thr Phe Glu Arg Asp Ser Lys  
 180 185 190  
 Ser Glu Lys Gly Thr Tyr Asp Leu Thr Leu Met Tyr Ile Leu Asp Ala  
 195 200 205  
 Tyr Arg Arg Asn Pro Lys Cys Gln Thr Ser Leu Ile Ala Ser Asp Ala  
 210 215 220  
 Asn Trp Met Ala Tyr Gly Thr Asn Leu Ala Arg Gln Gly Phe Ser Val  
 225 230 235 240  
 Arg Ala Ala Asn His Phe Gln Val Asn Pro His Met Glu Ile Phe Gly  
 245 250 255  
 Gln Phe Ala Phe Glu Val Arg Ser Ser Ser Arg Asn Tyr Asn Thr Asn  
 260 265 270  
 Leu Gly Ser Lys Phe Cys Phe  
 275

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1545 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATGACCATAC	TTCGAAATTT	TCTTACCTGC	TCGGCTTTAT	TCCTCGCTCT	CCCTGCAGCA	60
GCACAAGTTG	TATATCTTCA	TGAAAGTGAT	GGTTATAACG	GTGCTATCAA	TAATAAAAGC	120
TTAGAACCTA	AAATTACCTG	TTATCCAGAA	GGAACCTTCT	ACATCTTTCT	AGATGACGTG	180
AGGATTTCCA	ACGTTAAGCA	TGATCAAGAA	GATGCTGGGG	TTTTTATAAA	TCGATCTGGG	240
AATCTTTTTT	TCATGGGCAA	CCGTTGCAAC	TTCACTTTTT	ACAACCTTAT	GACCGAGGGT	300
TTTGGCGCTG	CCATTTTCGAA	CCGCGTTGGA	GACACCACTC	TCACTCTCTC	TAATTTTCT	360
TACTTAACGT	TCACCTCAGC	ACCTCTACTA	CCTCAAGGAC	AAGGAGCGAT	TTATAGTCTT	420
GGTTCGCTGA	TGATCGAAAA	TAGTGAGGAA	GTGACTTTCT	GTGGGAACTA	CTCTTCGTGG	480
AGTGGAGCTG	CGATTTTATAC	TCCCTACCTT	TTAGGTTCTA	AGGCGAGTCG	TCCTTCAGTA	540
AATCTCAGCG	GGAACCGCTA	CCTGGTGTTT	AGAGACTATG	TGAGCCAAGG	TTATGGCGGC	600
GCCGTATCTA	CCCACAATCT	CACACTCACG	ACTCGAGGAC	CTTCGTGTTT	TGAAAATAAT	660
CATGCTTATC	ATGACGTGAA	TAGTAATGGA	GGAGCCATTG	CCATTGCTCC	TGGAGGATCG	720



```

ATCTCTATAT CCGTGAAAAG CGGAGATCTC ATCTTCAAAG GAAATACAGC ATCACAAGAC      780
GGAAATACAA TACACAACCTC CATCCATCTG CAATCTGGAG CACAGTTTAA GAACCTACGT      840
GCTGTTTCAG AATCCGGAGT TTATTTCTAT GATCCTATAA GCCATAGCGA GTCGCATAAA      900
ATTACAGATC TTGTAATCAA TGCTCCTGAA GGAAAGGAAA CTTATGAAGG AACAAATTAGC      960
TTCTCAGGAC TATGCCTGGA TGATCATGAA GTTTGTGCGG AAAATCTTAC TTCCACAATC     1020
CTACAAGATG TCACATTAGC AGGAGGAACT CTCTCTCTAT CGGATGGGGT TACCTTGCAA     1080
CTGCATTCTT TTAAGCAGGA AGCAAGCTCT ACGCTTACTA TGTCTCCAGG AACCACCTCTG     1140
CTCTGCTCAG GAGATGCTCG GGTTCAGAAT CTGCACATCC TGATTGAAGA TACCGACAAC     1200
TTTGTTCTCG TAAGGATTCG CGCCGAGGAC AAGGATGCTC TTGTCTCATT AGAAAACTT      1260
AAAGTTGCCT TTGAGGCTTA TTGGTCCGTC TATGACTTTC CTCAATTTAA GGAAGCCTTT     1320
ACGATTCCCTC TTCTTGAAC TCTAGGGCCT TCTTTTGACA GTCTTCTCCT AGGGGAGACC     1380
ACTTTGGAGA GAACCCAAGT CACAACAGAG AATGACGCCG TTCGAGGTTT CTGGTCCCTA     1440
AGCTGGGAAG AGTACCCCCC TTCTCTGGAT AAAGACAGAA GGATCACACC AACTAAGAAA     1500
ACTGTTTTCC TCACTTGGA TCCTGAGATC ACTTCTACGC CATAA      1545

```

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 514 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```

Met Thr Ile Leu Arg Asn Phe Leu Thr Cys Ser Ala Leu Phe Leu Ala
 1             5             10             15

Leu Pro Ala Ala Ala Gln Val Val Tyr Leu His Glu Ser Asp Gly Tyr
      20             25             30

Asn Gly Ala Ile Asn Asn Lys Ser Leu Glu Pro Lys Ile Thr Cys Tyr
      35             40             45

Pro Glu Gly Thr Ser Tyr Ile Phe Leu Asp Asp Val Arg Ile Ser Asn
      50             55             60

Val Lys His Asp Gln Glu Asp Ala Gly Val Phe Ile Asn Arg Ser Gly
      65             70             75             80

Asn Leu Phe Phe Met Gly Asn Arg Cys Asn Phe Thr Phe His Asn Leu
      85             90             95

Met Thr Glu Gly Phe Gly Ala Ala Ile Ser Asn Arg Val Gly Asp Thr
      100            105            110

Thr Leu Thr Leu Ser Asn Phe Ser Tyr Leu Thr Phe Thr Ser Ala Pro
      115            120            125

Leu Leu Pro Gln Gly Gln Gly Ala Ile Tyr Ser Leu Gly Ser Val Met
      130            135            140

Ile Glu Asn Ser Glu Glu Val Thr Phe Cys Gly Asn Tyr Ser Ser Trp
      145            150            155            160

Ser Gly Ala Ala Ile Tyr Thr Pro Tyr Leu Leu Gly Ser Lys Ala Ser

```

165										170					175				
Arg	Pro	Ser	Val	Asn	Leu	Ser	Gly	Asn	Arg	Arg	Tyr	Leu	Val	Phe	Arg	Asp			
			180					185						190					
Tyr	Val	Ser	Gln	Gly	Tyr	Gly	Gly	Ala	Val	Ser	Thr	His	Asn	Leu	Thr				
		195				200						205							
Leu	Thr	Thr	Arg	Gly	Pro	Ser	Cys	Phe	Glu	Asn	Asn	His	Ala	Tyr	His				
	210					215					220								
Asp	Val	Asn	Ser	Asn	Gly	Gly	Ala	Ile	Ala	Ile	Ala	Pro	Gly	Gly	Ser				
225					230					235					240				
Ile	Ser	Ile	Ser	Val	Lys	Ser	Gly	Asp	Leu	Ile	Phe	Lys	Gly	Asn	Thr				
				245					250					255					
Ala	Ser	Gln	Asp	Gly	Asn	Thr	Ile	His	Asn	Ser	Ile	His	Leu	Gln	Ser				
			260					265					270						
Gly	Ala	Gln	Phe	Lys	Asn	Leu	Arg	Ala	Val	Ser	Glu	Ser	Gly	Val	Tyr				
		275					280					285							
Phe	Tyr	Asp	Pro	Ile	Ser	His	Ser	Glu	Ser	His	Lys	Ile	Thr	Asp	Leu				
	290					295					300								
Val	Ile	Asn	Ala	Pro	Glu	Gly	Lys	Glu	Thr	Tyr	Glu	Gly	Thr	Ile	Ser				
305					310					315					320				
Phe	Ser	Gly	Leu	Cys	Leu	Asp	Asp	His	Glu	Val	Cys	Ala	Glu	Asn	Leu				
				325					330					335					
Thr	Ser	Thr	Ile	Leu	Gln	Asp	Val	Thr	Leu	Ala	Gly	Gly	Thr	Leu	Ser				
			340					345					350						
Leu	Ser	Asp	Gly	Val	Thr	Leu	Gln	Leu	His	Ser	Phe	Lys	Gln	Glu	Ala				
		355					360					365							
Ser	Ser	Thr	Leu	Thr	Met	Ser	Pro	Gly	Thr	Thr	Leu	Leu	Cys	Ser	Gly				
	370					375					380								
Asp	Ala	Arg	Val	Gln	Asn	Leu	His	Ile	Leu	Ile	Glu	Asp	Thr	Asp	Asn				
385					390					395					400				
Phe	Val	Pro	Val	Arg	Ile	Arg	Ala	Glu	Asp	Lys	Asp	Ala	Leu	Val	Ser				
				405				410					415						
Leu	Glu	Lys	Leu	Lys	Val	Ala	Phe	Glu	Ala	Tyr	Trp	Ser	Val	Tyr	Asp				
			420					425					430						
Phe	Pro	Gln	Phe	Lys	Glu	Ala	Phe	Thr	Ile	Pro	Leu	Leu	Glu	Leu	Leu				
		435					440					445							
Gly	Pro	Ser	Phe	Asp	Ser	Leu	Leu	Leu	Gly	Glu	Thr	Thr	Leu	Glu	Arg				
						455					460								
Thr	Gln	Val	Thr	Thr	Glu	Asn	Asp	Ala	Val	Arg	Gly	Phe	Trp	Ser	Leu				
465					470					475					480				

Ser Trp Glu Glu Tyr Pro Pro Ser Leu Asp Lys Asp Arg Arg Ile Thr  
                     485                    490                    495

Pro Thr Lys Lys Thr Val Phe Leu Thr Trp Asn Pro Glu Ile Thr Ser  
                     500                    505                    510

Thr Pro

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 787 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ATGAAAACGT	CTATTCGTAA	GTTCTTAATT	TCTACCACAC	TGGCGCCATG	TTTTGCTTCA	60
ACAGCGTTTA	CTGTAGAAGT	TATCATGCCT	TCCGAGAACT	TTGATGGATC	GAGTGGGAAG	120
ATTTTTCCTT	ACACAACACT	TTCTGATCCT	AGAGGGACAC	TCTGTATTTT	TTCAGGGGAT	180
CTCTACATTG	CGAATCTTGA	TAATGCCATA	TCCAGAACCT	CTTCCAGTTG	CTTTAGCAAT	240
AGGGCGGGAG	CACTACAAAT	CTTAGGAAAA	GGTGGGGTTT	TCTCCTTCTT	AAATATCCGT	300
TCTTCAGCTG	ACGGAGCCGC	GATTAGTAGT	GTAATCACCC	AAAATCCTGA	ACTATGTCCC	360
TTGAGTTTTT	CAGGATTTAG	TCAGATGATC	TTCGATAACT	GTGAATCTTT	GACTTCAGAT	420
ACCTCAGCGA	GTAATGTCAT	ACCTCACGCA	TCGGCGATTT	ACGCTACAAC	GCCCATGCTC	480
TTTACAAACA	ATGACTCCAT	ACTATTCCAA	TACAACCGTT	CTGCAGGATT	TGGAGCTGCC	540
ATTCGAGGCA	CAAGCATCAC	AATAGAAAAT	ACGAAAAAGA	GCCTTCTCTT	TAATGGTAAT	600
GGATCCATCT	CTAATGGAGG	GGCCCTCACG	GGATCTGCAG	CGATCAACCT	CATCAACAAT	660
AGCGCTCCTG	TGATTTTCTC	AACGAATGCT	ACAGGGATCT	ATGGTGGGGC	TATTTACCTT	720
ACCGGAGGAT	CTATGCTCAC	CTCTGGGAAC	CTCTCAGGAG	TCTTGTTTCG	TTATAATAGC	780
TCGCGCT						787

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 262 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met	Lys	Thr	Ser	Ile	Arg	Lys	Phe	Leu	Ile	Ser	Thr	Thr	Leu	Ala	Pro
1				5				10					15		
Cys	Phe	Ala	Ser	Thr	Ala	Phe	Thr	Val	Glu	Val	Ile	Met	Pro	Ser	Glu
			20					25					30		
Asn	Phe	Asp	Gly	Ser	Ser	Gly	Lys	Ile	Phe	Pro	Tyr	Thr	Thr	Leu	Ser
			35				40					45			

Asp Pro Arg Gly Thr Leu Cys Ile Phe Ser Gly Asp Leu Tyr Ile Ala  
 50 55 60  
 Asn Leu Asp Asn Ala Ile Ser Arg Thr Ser Ser Ser Cys Phe Ser Asn  
 65 70 75 80  
 Arg Ala Gly Ala Leu Gln Ile Leu Gly Lys Gly Gly Val Phe Ser Phe  
 85 90 95  
 Leu Asn Ile Arg Ser Ser Ala Asp Gly Ala Ala Ile Ser Ser Val Ile  
 100 105 110  
 Thr Gln Asn Pro Glu Leu Cys Pro Leu Ser Phe Ser Gly Phe Ser Gln  
 115 120 125  
 Met Ile Phe Asp Asn Cys Glu Ser Leu Thr Ser Asp Thr Ser Ala Ser  
 130 135 140  
 Asn Val Ile Pro His Ala Ser Ala Ile Tyr Ala Thr Thr Pro Met Leu  
 145 150 155 160  
 Phe Thr Asn Asn Asp Ser Ile Leu Phe Gln Tyr Asn Arg Ser Ala Gly  
 165 170 175  
 Phe Gly Ala Ala Ile Arg Gly Thr Ser Ile Thr Ile Glu Asn Thr Lys  
 180 185 190  
 Lys Ser Leu Leu Phe Asn Gly Asn Gly Ser Ile Ser Asn Gly Gly Ala  
 195 200 205  
 Leu Thr Gly Ser Ala Ala Ile Asn Leu Ile Asn Asn Ser Ala Pro Val  
 210 215 220  
 Ile Phe Ser Thr Asn Ala Thr Gly Ile Tyr Gly Gly Ala Ile Tyr Leu  
 225 230 235 240  
 Thr Gly Gly Ser Met Leu Thr Ser Gly Asn Leu Ser Gly Val Leu Phe  
 245 250 255  
 Val Tyr Asn Ser Ser Arg  
 260

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2838 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: Genomic DNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ATGAAGACTT	CAGTTTCTAT	GTTGTTGGCC	CTGCTTTGCT	CGGGGGCTAG	CTCTATTGTA	60
CTCCATGCCG	CAACCACTCC	ACTAAATCCT	GAAGATGGGT	TTATTGGGGA	GGGCAATACA	120
AATACTTTTT	CTCCGAAATC	TACAACGGAT	GCTGCAGGAA	CTACCTACTC	TCTCACAGGA	180
GAGGTTCTGT	TTATAGATCC	GGGGAAAGGT	GTTTCAATTA	CAGGAAGTTG	CTTTGTAGAA	240

ACTGCTGGCG	ATCTTACATT	TTTAGGTAAT	GGAAATACCC	TAAAGTTCCT	GTCGGTAGAT	300
GCAGGTGCTA	ATATCGCGGT	TGCTCATGTA	CAAGGAAGTA	AGAATTTAAG	CTTCACAGAT	360
TTCCTTTCTC	TGGTGATCAC	AGAATCTCCA	AAATCCGCTG	TTAGTACAGG	AAAAGGTAGC	420
CTAGTCAGTT	CAGGTGCAGT	CCAACTGCAA	GATATAAACA	CTCTAGTTCT	TACAAGCAAT	480
GCCTCTGTCT	AAGATGGTGG	CGTGATTAAA	GGAAACTCCT	GCTTGATTCA	GGGAATCAAA	540
AATAGTGCGA	TTTTTGGACA	AAATACATCT	TCGAAAAAAG	GAGGGGCGAT	CTCCACGACT	600
CAAGGACTCA	CCATAGAGAA	TAACCTAGGG	ACGCTAAAGT	TCAATGAAAA	CAAAGCAGTG	660
ACCTCAGGAG	GCGCCTTAGA	TTTAGGAGCC	GCGTCTACAT	TCACTGCGAA	CCATGAGTTG	720
ATATTTTTCAC	AAAATAAGAC	TTCTGGGAAT	GCTGCAAATG	GCGGAGCCAT	AAATTGCTCA	780
GGCGACCTAA	CATTTACTGA	TAACACTTCT	TTGTTACTTC	AAGAAAATAG	CACAATGCAG	840
GATGGTGGAG	CTTTGTGTAG	CACAGGAACC	ATAAGCATTA	CCGGTAGTGA	TTCTATCAAT	900
GTGATAGGAA	ATACTTCAGG	ACAAAAAGGA	GGAGCGATTT	CTGCAGCTTC	TCTCAAGATT	960
TTGGGAGGGC	AGGGAGGCGC	TCTCTTTTCT	AATAACGTAG	TGACTCATGC	CACCCCTCTA	1020
GGAGGTGCCA	TTTTTATCAA	CACAGGAGGA	TCCTTGCGAGC	TCTTCACTCA	AGGAGGGGGAT	1080
ATCGTATTCG	AGGGGAATCA	GGTCACTACA	ACAGCTCCAA	ATGCTACCAC	TAAGAGAAAT	1140
GTAATTCACC	TCGAGAGCAC	CGCGAAGTGG	ACGGGACTTG	CTGCAAGTCA	AGGTAACGCT	1200
ATCTATTTCT	TACGATCCCAT	TACCACCAAC	GATACGGGAG	CAAGCGATAA	CTTACGTATC	1260
AATGAGGTCA	GTGCAAATCA	AAAGCTCTCG	GGATCTATAG	TATTTTCTGG	AGAGAGATTG	1320
TCGACAGCAG	AAGCTATAGC	TGAAAACTCT	ACTTCGAGGA	TCAACCAGCC	TGTCACTTTA	1380
GTAGAGGGGA	GCTTAGAACT	TAAACAGGGA	GTGACCTTGA	TCACACAAGG	ATTCTCGCAG	1440
GAGCCAGAAT	CCACGCTTCT	TTTGGATTG	GGGACCTCAT	TACAAGCTTC	TACAGAAGAT	1500
ATCGTCATCA	CAAATTCATC	TATAAATGCC	GATACCATTT	ACGGAAAGAA	TCCAATCAAT	1560
ATTGTAGCTT	CAGCAGCGAA	TAAGAACATT	ACCCTAACAG	GAACCTTAGC	ACTTGTAAT	1620
GCAGATGGAG	CTTTGTATGA	GAACCATACC	TTGCAAGACT	CTCAAGATTA	TAGCTTTGTA	1680
AAGTTATCTC	CAGGAGCGGG	AGGGACTATA	ATTACTCAAG	ATGCTTCTCA	GAAGCTTCTT	1740
GAAGTAGCTC	CTTCTAGACC	ACATTATGGC	TATCAAGGAC	ATTGGAATGT	GCAAGTCATC	1800
CCAGGAACGG	GAACTCAACC	GAGCCAGGCA	AATTTAGAAT	GGGTGCGGAC	AGGATACCTT	1860
CCGAATCCCG	AACGGCAAGG	ATTTTGTAGT	CCCAATAGCC	TGTGGGGTTC	TTTTGTGAT	1920
CAGCGTGCTA	TCCAAGAAAT	CATGGTAAAT	AGTAGCCAAA	TCTTATGTCA	GGAACGGGGA	1980
GTCTGGGGAG	CTGGAATTGC	TAATTTCCCTA	CATAGAGATA	AAATTAATGA	GCACGGCTAT	2040
CGCCATAGCG	GTGTGGGTTA	TCTTGTGGGA	GTTGGCACTC	ATGCTTTTTT	TGATGCTACG	2100
ATAAATGCGG	CTTTTTGCCA	GCTCTTCAGT	AGAGATAAAG	ACTACGTAGT	ATCCAAAAAT	2160
CATGGAACCTA	GCTACTCAGG	GGTCGTATTT	CTTGAGGATA	CCCTAGAGTT	TAGAAGTCCA	2220
CAGGATTCTC	ATACTGATAG	CTCCTCAGAA	GCTTGCTGTA	ACCAAGTCGT	CACTATAGAT	2280
ATGCAGTTGT	CTTACAGCCA	TAGAAAATAAT	GATATGAAAA	CCAAATACAC	GACATATCCA	2340
GAAGCTCAGG	GATCTTGGGC	AAATGATGTT	TTTGGTCTTG	AGTTTGGAGC	GACTACATAC	2400
TACTACCCTA	ACAGTACTTT	TTTATTTGAT	TACTACTCTC	CGTTTCTCAG	GCTGCAGTGC	2460
ACCTATGCTC	ACCAGGAAGA	CTTCAAAGAG	ACAGGAGGTG	AGGTTCGTCA	CTTTACTAGC	2520
GGAGATCTTT	TCAATTTAGC	AGTTCCTATT	GGCGTGAAGT	TTGAGAGATT	TTCAGACTGT	2580
AAAAGGGGAT	CTTATGAACT	TACCCTTGCT	TATGTTCCCTG	ATGTGATTCTG	CAAAGATCCC	2640
AAGAGCACGG	CAACATTGGC	TAGTGGAGCT	ACGTGGAGCA	CCCACGGAAA	CAATCTCTCC	2700
AGACAAGGAT	TACAACTGCG	TTTAGGGGAA	CACTGTCTCA	TAAATCCTGG	AATTGAGGTG	2760
TTCAGTCAG	GAGCTATTGA	ATTGCGGGGA	TCCTCTCGTA	ATTATAACAT	CAATCTCGGG	2820
GGTAAATACC	GATTTTAA					2838

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 946 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Lys Thr Ser Val Ser Met Leu Leu Ala Leu Leu Cys Ser Gly Ala

1	5	10	15
Ser Ser Ile Val	Leu His Ala Ala	Thr Thr Pro Leu Asn	Pro Glu Asp
20		25	30
Gly Phe Ile Gly	Glu Gly Asn Thr	Asn Thr Phe Ser	Pro Lys Ser Thr
35	40	45	
Thr Asp Ala Ala	Gly Thr Thr Tyr	Ser Leu Thr Gly	Glu Val Leu Phe
50	55	60	
Ile Asp Pro Gly	Lys Gly Gly Ser	Ile Thr Gly Thr	Cys Phe Val Glu
65	70	75	80
Thr Ala Gly Asp	Leu Thr Phe Leu	Gly Asn Gly Asn	Thr Leu Lys Phe
	85	90	95
Leu Ser Val Asp	Ala Gly Ala Asn	Ile Ala Val Ala	His Val Gln Gly
100		105	110
Ser Lys Asn Leu	Ser Phe Thr Asp	Phe Leu Ser Leu	Val Ile Thr Glu
115	120	125	
Ser Pro Lys Ser	Ala Val Ser Thr	Gly Lys Gly Ser	Leu Val Ser Ser
130	135	140	
Gly Ala Val Gln	Leu Gln Asp Ile	Asn Thr Leu Val	Leu Thr Ser Asn
145	150	155	160
Ala Ser Val Glu	Asp Gly Gly Val	Ile Lys Gly Asn	Ser Cys Leu Ile
165		170	175
Gln Gly Ile Lys	Asn Ser Ala Ile	Phe Gly Gln Asn	Thr Ser Ser Lys
180		185	190
Lys Gly Gly Ala	Ile Ser Thr Thr	Gln Gly Leu Thr	Ile Glu Asn Asn
195	200	205	
Leu Gly Thr Leu	Lys Phe Asn Glu	Asn Lys Ala Val	Thr Ser Gly Gly
210	215	220	
Ala Leu Asp Leu	Gly Ala Ala Ser	Thr Phe Thr Ala	Asn His Glu Leu
225	230	235	240
Ile Phe Ser Gln	Asn Lys Thr Ser	Gly Asn Ala Ala	Asn Gly Gly Ala
	245	250	255
Ile Asn Cys Ser	Gly Asp Leu Thr	Phe Thr Asp Asn	Thr Ser Leu Leu
260		265	270
Leu Gln Glu Asn	Ser Thr Met Gln	Asp Gly Gly Ala	Leu Cys Ser Thr
275	280	285	
Gly Thr Ile Ser	Ile Thr Gly Ser	Asp Ser Ile Asn	Val Ile Gly Asn
290	295	300	
Thr Ser Gly Gln	Lys Gly Gly Ala	Ile Ser Ala Ala	Ser Leu Lys Ile
305	310	315	320

Leu Gly Gly Gln Gly Gly Ala Leu Phe Ser Asn Asn Val Val Thr His  
 325 330 335  
 Ala Thr Pro Leu Gly Gly Ala Ile Phe Ile Asn Thr Gly Gly Ser Leu  
 340 345 350  
 Gln Leu Phe Thr Gln Gly Gly Asp Ile Val Phe Glu Gly Asn Gln Val  
 355 360 365  
 Thr Thr Thr Ala Pro Asn Ala Thr Thr Lys Arg Asn Val Ile His Leu  
 370 375 380  
 Glu Ser Thr Ala Lys Trp Thr Gly Leu Ala Ala Ser Gln Gly Asn Ala  
 385 390 395 400  
 Ile Tyr Phe Tyr Asp Pro Ile Thr Thr Asn Asp Thr Gly Ala Ser Asp  
 405 410 415  
 Asn Leu Arg Ile Asn Glu Val Ser Ala Asn Gln Lys Leu Ser Gly Ser  
 420 425 430  
 Ile Val Phe Ser Gly Glu Arg Leu Ser Thr Ala Glu Ala Ile Ala Glu  
 435 440 445  
 Asn Leu Thr Ser Arg Ile Asn Gln Pro Val Thr Leu Val Glu Gly Ser  
 450 455 460  
 Leu Glu Leu Lys Gln Gly Val Thr Leu Ile Thr Gln Gly Phe Ser Gln  
 465 470 475 480  
 Glu Pro Glu Ser Thr Leu Leu Leu Asp Leu Gly Thr Ser Leu Gln Ala  
 485 490 495  
 Ser Thr Glu Asp Ile Val Ile Thr Asn Ser Ser Ile Asn Ala Asp Thr  
 500 505 510  
 Ile Tyr Gly Lys Asn Pro Ile Asn Ile Val Ala Ser Ala Ala Asn Lys  
 515 520 525  
 Asn Ile Thr Leu Thr Gly Thr Leu Ala Leu Val Asn Ala Asp Gly Ala  
 530 535 540  
 Leu Tyr Glu Asn His Thr Leu Gln Asp Ser Gln Asp Tyr Ser Phe Val  
 545 550 555 560  
 Lys Leu Ser Pro Gly Ala Gly Gly Thr Ile Ile Thr Gln Asp Ala Ser  
 565 570 575  
 Gln Lys Leu Leu Glu Val Ala Pro Ser Arg Pro His Tyr Gly Tyr Gln  
 580 585 590  
 Gly His Trp Asn Val Gln Val Ile Pro Gly Thr Gly Thr Gln Pro Ser  
 595 600 605  
 Gln Ala Asn Leu Glu Trp Val Arg Thr Gly Tyr Leu Pro Asn Pro Glu  
 610 615 620

Arg Gln Gly Phe Leu Val Pro Asn Ser Leu Trp Gly Ser Phe Val Asp  
 625 630 635 640  
 Gln Arg Ala Ile Gln Glu Ile Met Val Asn Ser Ser Gln Ile Leu Cys  
 645 650 655  
 Gln Glu Arg Gly Val Trp Gly Ala Gly Ile Ala Asn Phe Leu His Arg  
 660 665 670  
 Asp Lys Ile Asn Glu His Gly Tyr Arg His Ser Gly Val Gly Tyr Leu  
 675 680 685  
 Val Gly Val Gly Thr His Ala Phe Ser Asp Ala Thr Ile Asn Ala Ala  
 690 695 700  
 Phe Cys Gln Leu Phe Ser Arg Asp Lys Asp Tyr Val Val Ser Lys Asn  
 705 710 715 720  
 His Gly Thr Ser Tyr Ser Gly Val Val Phe Leu Glu Asp Thr Leu Glu  
 725 730 735  
 Phe Arg Ser Pro Gln Gly Phe Tyr Thr Asp Ser Ser Ser Glu Ala Cys  
 740 745 750  
 Cys Asn Gln Val Val Thr Ile Asp Met Gln Leu Ser Tyr Ser His Arg  
 755 760 765  
 Asn Asn Asp Met Lys Thr Lys Tyr Thr Thr Tyr Pro Glu Ala Gln Gly  
 770 775 780  
 Ser Trp Ala Asn Asp Val Phe Gly Leu Glu Phe Gly Ala Thr Thr Tyr  
 785 790 795 800  
 Tyr Tyr Pro Asn Ser Thr Phe Leu Phe Asp Tyr Tyr Ser Pro Phe Leu  
 805 810 815  
 Arg Leu Gln Cys Thr Tyr Ala His Gln Glu Asp Phe Lys Glu Thr Gly  
 820 825 830  
 Gly Glu Val Arg His Phe Thr Ser Gly Asp Leu Phe Asn Leu Ala Val  
 835 840 845  
 Pro Ile Gly Val Lys Phe Glu Arg Phe Ser Asp Cys Lys Arg Gly Ser  
 850 855 860  
 Tyr Glu Leu Thr Leu Ala Tyr Val Pro Asp Val Ile Arg Lys Asp Pro  
 865 870 875 880  
 Lys Ser Thr Ala Thr Leu Ala Ser Gly Ala Thr Trp Ser Thr His Gly  
 885 890 895  
 Asn Asn Leu Ser Arg Gln Gly Leu Gln Leu Arg Leu Gly Asn His Cys  
 900 905 910  
 Leu Ile Asn Pro Gly Ile Glu Val Phe Ser His Gly Ala Ile Glu Leu  
 915 920 925  
 Arg Gly Ser Ser Arg Asn Tyr Asn Ile Asn Leu Gly Gly Lys Tyr Arg



930

935

940

Phe  
945

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3000 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 259...3000
- (D) OTHER INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ATCAGGTGAT AAAAGTTCCT CGTTAGCTAG TGA	CTGTAGG TGACATGAGA AAGCTAACAC	60
GGAGGAAACT AAAACCCAAG GAATCGAAGT C	TTTCATGGTA ATGCTTTTGT TTTT	120
ACTATTCGCA TCAATATAGA AACAAAATAA G	TAAATCAAG TTAAAGATGA CAAAACAGCT	180
GTCAAGAATT TTTATCTTGA CTCTCTGAGT T	TTCTATTTT ATATGACGCA AGTAAGAATT	240
TAATAATAAA GTGGGTTT ATG AAA TCG CAA	TTT TCC TGG TTA GTG CTC TCT	291
Met Lys Ser Gln Phe Ser Trp Leu Val Leu Ser		
1 5 10		
TCG ACA TTG GCA TGT TTT ACT AGT TGT	TCC ACT GTT TTT GCT GCA ACT	339
Ser Thr Leu Ala Cys Phe Thr Ser Cys	Ser Thr Val Phe Ala Ala Thr	
15 20 25		
GCT GAA AAT ATA GGC CCC TCT GAT AGC	TTT GAC GGA AGT ACT AAC ACA	387
Ala Glu Asn Ile Gly Pro Ser Asp Ser	Phe Asp Gly Ser Thr Asn Thr	
30 35 40		
GGC ACC TAT ACT CCT AAA AAT ACG ACT	ACT GGA ATA GAC TAT ACT CTG	435
Gly Thr Tyr Thr Pro Lys Asn Thr Thr	Gly Ile Asp Tyr Thr Leu	
45 50 55		
ACA GGA GAT ATA ACT CTG CAA AAC CTT	GGG GAT TCG GCA GCT TTA ACG	483
Thr Gly Asp Ile Thr Leu Gln Asn Leu	Gly Asp Ser Ala Ala Leu Thr	
60 65 70 75		
AAG GGT TGT TTT TCT GAC ACT ACG GAA	TCT TTA AGC TTT GCC GGT AAG	531
Lys Gly Cys Phe Ser Asp Thr Thr Glu	Ser Leu Ser Phe Ala Gly Lys	
80 85 90		
GGG TAC TCA CTT TCT TTT TTA AAT ATT	AAG TCT AGT GCT GAA GGC GCA	579
Gly Tyr Ser Leu Ser Phe Leu Asn Ile	Lys Ser Ser Ala Glu Gly Ala	
95 100 105		
GCA CTT TCT GTT ACA ACT GAT AAA AAT	CTG TCG CTA ACA GGA TTT TCG	627

Ala	Leu	Ser	Val	Thr	Thr	Asp	Lys	Asn	Leu	Ser	Leu	Thr	Gly	Phe	Ser	
	110						115					120				
AGT	CTT	ACT	TTC	TTA	GCG	GCC	CCA	TCA	TCG	GTA	ATC	ACA	ACC	CCC	TCA	675
Ser	Leu	Thr	Phe	Leu	Ala	Ala	Pro	Ser	Ser	Val	Ile	Thr	Thr	Pro	Ser	
	125					130					135					
GGA	AAA	GGT	GCA	GTT	AAA	TGT	GGA	GGG	GAT	CTT	ACA	TTT	GAT	AAC	AAT	723
Gly	Lys	Gly	Ala	Val	Lys	Cys	Gly	Gly	Asp	Leu	Thr	Phe	Asp	Asn	Asn	
140					145					150					155	
GGA	ACT	ATT	TTA	TTT	AAA	CAA	GAT	TAC	TGT	GAG	GAA	AAT	GGC	GGA	GCC	771
Gly	Thr	Ile	Leu	Phe	Lys	Gln	Asp	Tyr	Cys	Glu	Glu	Asn	Gly	Gly	Ala	
				160					165						170	
ATT	TCT	ACC	AAG	AAT	CTT	TCT	TTG	AAA	AAC	AGC	ACG	GGA	TCG	ATT	TCT	819
Ile	Ser	Thr	Lys	Asn	Leu	Ser	Leu	Lys	Asn	Ser	Thr	Gly	Ser	Ile	Ser	
			175					180					185			
TTT	GAA	GGG	AAT	AAA	TCG	AGC	GCA	ACA	GGG	AAA	AAA	GGT	GGG	GCT	ATT	867
Phe	Glu	Gly	Asn	Lys	Ser	Ser	Ala	Thr	Gly	Lys	Lys	Gly	Gly	Ala	Ile	
	190						195					200				
TGT	GCT	ACT	GGT	ACT	GTA	GAT	ATT	ACA	AAT	AAT	ACG	GCT	CCT	ACC	CTC	915
Cys	Ala	Thr	Gly	Thr	Val	Asp	Ile	Thr	Asn	Asn	Thr	Ala	Pro	Thr	Leu	
	205					210					215					
TTC	TCG	AAC	AAT	ATT	GCT	GAA	GCT	GCA	GGT	GGA	GCT	ATA	AAT	AGC	ACA	963
Phe	Ser	Asn	Asn	Ile	Ala	Glu	Ala	Ala	Gly	Gly	Ala	Ile	Asn	Ser	Thr	
220					225					230					235	
GGA	AAC	TGT	ACA	ATT	ACA	GGG	AAT	ACG	TCT	CTT	GTA	TTT	TCT	GAA	AAT	1011
Gly	Asn	Cys	Thr	Ile	Thr	Gly	Asn	Thr	Ser	Leu	Val	Phe	Ser	Glu	Asn	
				240					245					250		
AGT	GTG	ACA	GCG	ACC	GCA	GGA	AAT	GGA	GGA	GCT	CTT	TCT	GGA	GAT	GCC	1059
Ser	Val	Thr	Ala	Thr	Ala	Gly	Asn	Gly	Gly	Ala	Leu	Ser	Gly	Asp	Ala	
			255					260					265			
GAT	GTT	ACC	ATA	TCT	GGG	AAT	CAG	AGT	GTA	ACT	TTC	TCA	GGA	AAC	CAA	1107
Asp	Val	Thr	Ile	Ser	Gly	Asn	Gln	Ser	Val	Thr	Phe	Ser	Gly	Asn	Gln	
		270					275					280				
GCT	GTA	GCT	AAT	GGC	GGA	GCC	ATT	TAT	GCT	AAG	AAG	CTT	ACA	CTG	GCT	1155
Ala	Val	Ala	Asn	Gly	Gly	Ala	Ile	Tyr	Ala	Lys	Lys	Leu	Thr	Leu	Ala	
	285					290					295					
TCC	GGG	GGG	GGG	GGG	GGT	ATC	TCC	TTT	TCT	AAC	AAT	ATA	GTC	CAA	GGT	1203
Ser	Gly	Gly	Gly	Gly	Gly	Ile	Ser	Phe	Ser	Asn	Asn	Ile	Val	Gln	Gly	
300					305					310					315	
ACC	ACT	GCA	GGT	AAT	GGT	GGA	GCC	ATT	TCT	ATA	CTG	GCA	GCT	GGA	GAG	1251
Thr	Thr	Ala	Gly	Asn	Gly	Gly	Ala	Ile	Ser	Ile	Leu	Ala	Ala	Gly	Glu	
				320					325					330		
TGT	AGT	CTT	TCA	GCA	GAA	GCA	GGG	GAC	ATT	ACC	TTC	AAT	GGG	AAT	GCC	1299
Cys	Ser	Leu	Ser	Ala	Glu	Ala	Gly	Asp	Ile	Thr	Phe	Asn	Gly	Asn	Ala	
			335					340					345			

ATT	GTT	GCA	ACT	ACA	CCA	CAA	ACT	ACA	AAA	AGA	AAT	TCT	ATT	GAC	ATA	1347
Ile	Val	Ala	Thr	Thr	Pro	Gln	Thr	Thr	Lys	Arg	Asn	Ser	Ile	Asp	Ile	
		350					355					360				
GGA	TCT	ACT	GCA	AAG	ATC	ACG	AAT	TTA	CGT	GCA	ATA	TCT	GGG	CAT	AGC	1395
Gly	Ser	Thr	Ala	Lys	Ile	Thr	Asn	Leu	Arg	Ala	Ile	Ser	Gly	His	Ser	
	365					370					375					
ATC	TTT	TTC	TAC	GAT	CCG	ATT	ACT	GCT	AAT	ACG	GCT	GCG	GAT	TCT	ACA	1443
Ile	Phe	Phe	Tyr	Asp	Pro	Ile	Thr	Ala	Asn	Thr	Ala	Ala	Asp	Ser	Thr	
380					385					390					395	
GAT	ACT	TTA	AAT	CTC	AAT	AAG	GCT	GAT	GCA	GGT	AAT	AGT	ACA	GAT	TAT	1491
Asp	Thr	Leu	Asn	Leu	Asn	Lys	Ala	Asp	Ala	Gly	Asn	Ser	Thr	Asp	Tyr	
				400					405					410		
AGT	GGG	TCG	ATT	GTT	TTT	TCT	GGT	GAA	AAG	CTC	TCT	GAA	GAT	GAA	GCA	1539
Ser	Gly	Ser	Ile	Val	Phe	Ser	Gly	Glu	Lys	Leu	Ser	Glu	Asp	Glu	Ala	
			415					420					425			
AAA	GTT	GCA	GAC	AAC	CTC	ACT	TCT	ACG	CTG	AAG	CAG	CCT	GTA	ACT	CTA	1587
Lys	Val	Ala	Asp	Asn	Leu	Thr	Ser	Thr	Leu	Lys	Gln	Pro	Val	Thr	Leu	
		430					435					440				
ACT	GCA	GGA	AAT	TTA	GTA	CTT	AAA	CGT	GGT	GTC	ACT	CTC	GAT	ACG	AAA	1635
Thr	Ala	Gly	Asn	Leu	Val	Leu	Lys	Arg	Gly	Val	Thr	Leu	Asp	Thr	Lys	
	445					450					455					
GGC	TTT	ACT	CAG	ACC	GCG	GGT	TCC	TCT	GTT	ATT	ATG	GAT	GCG	GGC	ACA	1683
Gly	Phe	Thr	Gln	Thr	Ala	Gly	Ser	Ser	Val	Ile	Met	Asp	Ala	Gly	Thr	
460					465				470						475	
ACG	TTA	AAA	GCA	AGT	ACA	GAG	GAG	GTC	ACT	TTA	ACA	GGT	CTT	TCC	ATT	1731
Thr	Leu	Lys	Ala	Ser	Thr	Glu	Glu	Val	Thr	Leu	Thr	Gly	Leu	Ser	Ile	
				480				485						490		
CCT	GTA	GAC	TCT	TTA	GGC	GAG	GGT	AAG	AAA	GTT	GTA	ATT	GCT	GCT	TCT	1779
Pro	Val	Asp	Ser	Leu	Gly	Glu	Gly	Lys	Lys	Val	Val	Ile	Ala	Ala	Ser	
			495					500					505			
GCA	GCA	AGT	AAA	AAT	GTA	GCC	CTT	AGT	GGT	CCG	ATT	CTT	CTT	TTG	GAT	1827
Ala	Ala	Ser	Lys	Asn	Val	Ala	Leu	Ser	Gly	Pro	Ile	Leu	Leu	Leu	Asp	
		510					515					520				
AAC	CAA	GGG	AAT	GCT	TAT	GAA	AAT	CAC	GAC	TTA	GGA	AAA	ACT	CAA	GAC	1875
Asn	Gln	Gly	Asn	Ala	Tyr	Glu	Asn	His	Asp	Leu	Gly	Lys	Thr	Gln	Asp	
	525					530					535					
TTT	TCA	TTT	GTG	CAG	CTC	TCT	GCT	CTG	GGT	ACT	GCA	ACA	ACT	ACA	GAT	1923
Phe	Ser	Phe	Val	Gln	Leu	Ser	Ala	Leu	Gly	Thr	Ala	Thr	Thr	Thr	Asp	
540					545				550						555	
GTT	CCA	GCG	GTT	CCT	ACA	GTA	GCA	ACT	CCT	ACG	CAC	TAT	GGG	TAT	CAA	1971
Val	Pro	Ala	Val	Pro	Thr	Val	Ala	Thr	Pro	Thr	His	Tyr	Gly	Tyr	Gln	
				560					565					570		
GGT	ACT	TGG	GGA	ATG	ACT	TGG	GTT	GAT	GAT	ACC	GCA	AGC	ACT	CCA	AAG	2019

Gly	Thr	Trp	Gly	Met	Thr	Trp	Val	Asp	Asp	Thr	Ala	Ser	Thr	Pro	Lys	
			575					580					585			
ACT	AAG	ACA	GCG	ACA	TTA	GCT	TGG	ACC	AAT	ACA	GGC	TAC	CTT	CCG	AAT	2067
Thr	Lys	Thr	Ala	Thr	Leu	Ala	Trp	Thr	Asn	Thr	Gly	Tyr	Leu	Pro	Asn	
			590				595					600				
CCT	GAG	CGT	CAA	GGA	CCT	TTA	GTT	CCT	AAT	AGC	CTT	TGG	GGA	TCT	TTT	2115
Pro	Glu	Arg	Gln	Gly	Pro	Leu	Val	Pro	Asn	Ser	Leu	Trp	Gly	Ser	Phe	
	605					610					615					
TCA	GAC	ATC	CAA	GCG	ATT	CAA	GGT	GTC	ATA	GAG	AGA	AGT	GCT	TTG	ACT	2163
Ser	Asp	Ile	Gln	Ala	Ile	Gln	Gly	Val	Ile	Glu	Arg	Ser	Ala	Leu	Thr	
620					625				630						635	
CTT	TGT	TCA	GAT	CGA	GGC	TTC	TGG	GCT	GCG	GGA	GTC	GCC	AAT	TTC	TTA	2211
Leu	Cys	Ser	Asp	Arg	Gly	Phe	Trp	Ala	Ala	Gly	Val	Ala	Asn	Phe	Leu	
				640					645					650		
GAT	AAA	GAT	AAG	AAA	GGG	GAA	AAA	CGC	AAA	TAC	CGT	CAT	AAA	TCT	GGT	2259
Asp	Lys	Asp	Lys	Lys	Gly	Glu	Lys	Arg	Lys	Tyr	Arg	His	Lys	Ser	Gly	
			655					660					665			
GGA	TAT	GCT	ATC	GGA	GGT	GCA	GCG	CAA	ACT	TGT	TCT	GAA	AAC	TTA	ATT	2307
Gly	Tyr	Ala	Ile	Gly	Gly	Ala	Ala	Gln	Thr	Cys	Ser	Glu	Asn	Leu	Ile	
		670					675					680				
AGC	TTT	GCC	TTT	TGC	CAA	CTC	TTT	GGT	AGC	GAT	AAA	GAT	TTC	TTA	GTC	2355
Ser	Phe	Ala	Phe	Cys	Gln	Leu	Phe	Gly	Ser	Asp	Lys	Asp	Phe	Leu	Val	
	685					690					695					
GCT	AAA	AAT	CAT	ACT	GAT	ACC	TAT	GCA	GGA	GCC	TTC	TAT	ATC	CAA	CAC	2403
Ala	Lys	Asn	His	Thr	Asp	Thr	Tyr	Ala	Gly	Ala	Phe	Tyr	Ile	Gln	His	
700					705				710						715	
ATT	ACA	GAA	TGT	AGT	GGG	TTC	ATA	GGT	TGT	CTC	TTA	GAT	AAA	CTT	CCT	2451
Ile	Thr	Glu	Cys	Ser	Gly	Phe	Ile	Gly	Cys	Leu	Leu	Asp	Lys	Leu	Pro	
				720				725						730		
GGC	TCT	TGG	AGT	CAT	AAA	CCC	CTC	GTT	TTA	GAA	GGG	CAG	CTC	GCT	TAT	2499
Gly	Ser	Trp	Ser	His	Lys	Pro	Leu	Val	Leu	Glu	Gly	Gln	Leu	Ala	Tyr	
			735					740					745			
AGC	CAC	GTC	AGT	AAT	GAT	CTG	AAG	ACA	AAG	TAT	ACT	GCG	TAT	CCT	GAG	2547
Ser	His	Val	Ser	Asn	Asp	Leu	Lys	Thr	Lys	Tyr	Thr	Ala	Tyr	Pro	Glu	
		750					755					760				
GTG	AAA	GGT	TCT	TGG	GGG	AAT	AAT	GCT	TTT	AAC	ATG	ATG	TTG	GGA	GCT	2595
Val	Lys	Gly	Ser	Trp	Gly	Asn	Asn	Ala	Phe	Asn	Met	Met	Leu	Gly	Ala	
	765					770				775						
TCT	TCT	CAT	TCT	TAT	CCT	GAA	TAC	CTG	CAT	TGT	TTT	GAT	ACC	TAT	GCT	2643
Ser	Ser	His	Ser	Tyr	Pro	Glu	Tyr	Leu	His	Cys	Phe	Asp	Thr	Tyr	Ala	
780					785					790					795	
CCA	TAC	ATC	AAA	CTG	AAT	CTG	ACC	TAT	ATA	CGT	CAG	GAC	AGC	TTC	TCG	2691
Pro	Tyr	Ile	Lys	Leu	Asn	Leu	Thr	Tyr	Ile	Arg	Gln	Asp	Ser	Phe	Ser	
				800					805					810		

GAG AAA GGT ACA GAA GGA AGA TCT TTT GAT GAC AGC AAC CTC TTC AAT	2739
Glu Lys Gly Thr Glu Gly Arg Ser Phe Asp Asp Ser Asn Leu Phe Asn	
815 820 825	
TTA TCT TTG CCT ATA GGG GTG AAG TTT GAG AAG TTC TCT GAT TGT AAT	2787
Leu Ser Leu Pro Ile Gly Val Lys Phe Glu Lys Phe Ser Asp Cys Asn	
830 835 840	
GAC TTT TCT TAT GAT CTG ACT TTA TCC TAT GTT CCT GAT CTT ATC CGC	2835
Asp Phe Ser Tyr Asp Leu Thr Leu Ser Tyr Val Pro Asp Leu Ile Arg	
845 850 855	
AAT GAT CCC AAA TGC ACT ACA GCA CTT GTA ATC AGC GGA GCC TCT TGG	2883
Asn Asp Pro Lys Cys Thr Thr Ala Leu Val Ile Ser Gly Ala Ser Trp	
860 865 870 875	
GAA ACT TAT GCC AAT AAC TTA GCA CGA CAG GCC TTG CAA GTG CGT GCA	2931
Glu Thr Tyr Ala Asn Asn Leu Ala Arg Gln Ala Leu Gln Val Arg Ala	
880 885 890	
GGC AGT CAC TAC GCC TTC TCT CCT ATG TTT GAA GTG CTC GGC CAG TTT	2979
Gly Ser His Tyr Ala Phe Ser Pro Met Phe Glu Val Leu Gly Gln Phe	
895 900 905	
GTC TTT GAA GTT CGT GGA TCC	3000
Val Phe Glu Val Arg Gly Ser	
910	

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 914 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (v) FRAGMENT TYPE: internal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Lys Ser Gln Phe Ser Trp Leu Val Leu Ser Ser Thr Leu Ala Cys	
1 5 10 15	
Phe Thr Ser Cys Ser Thr Val Phe Ala Ala Thr Ala Glu Asn Ile Gly	
20 25 30	
Pro Ser Asp Ser Phe Asp Gly Ser Thr Asn Thr Gly Thr Tyr Thr Pro	
35 40 45	
Lys Asn Thr Thr Thr Gly Ile Asp Tyr Thr Leu Thr Gly Asp Ile Thr	
50 55 60	
Leu Gln Asn Leu Gly Asp Ser Ala Ala Leu Thr Lys Gly Cys Phe Ser	
65 70 75 80	

Asp	Thr	Thr	Glu	Ser	Leu	Ser	Phe	Ala	Gly	Lys	Gly	Tyr	Ser	Leu	Ser		
				85					90					95			
Phe	Leu	Asn	Ile	Lys	Ser	Ser	Ala	Glu	Gly	Ala	Ala	Leu	Ser	Val	Thr		
			100					105					110				
Thr	Asp	Lys	Asn	Leu	Ser	Leu	Thr	Gly	Phe	Ser	Ser	Leu	Thr	Phe	Leu		
		115					120					125					
Ala	Ala	Pro	Ser	Ser	Val	Ile	Thr	Thr	Pro	Ser	Gly	Lys	Gly	Ala	Val		
	130					135					140						
Lys	Cys	Gly	Gly	Asp	Leu	Thr	Phe	Asp	Asn	Asn	Gly	Thr	Ile	Leu	Phe		
145					150				155					160			
Lys	Gln	Asp	Tyr	Cys	Glu	Glu	Asn	Gly	Gly	Ala	Ile	Ser	Thr	Lys	Asn		
			165					170						175			
Leu	Ser	Leu	Lys	Asn	Ser	Thr	Gly	Ser	Ile	Ser	Phe	Glu	Gly	Asn	Lys		
			180					185					190				
Ser	Ser	Ala	Thr	Gly	Lys	Lys	Gly	Gly	Ala	Ile	Cys	Ala	Thr	Gly	Thr		
		195					200					205					
Val	Asp	Ile	Thr	Asn	Asn	Thr	Ala	Pro	Thr	Leu	Phe	Ser	Asn	Asn	Ile		
	210					215					220						
Ala	Glu	Ala	Ala	Gly	Gly	Ala	Ile	Asn	Ser	Thr	Gly	Asn	Cys	Thr	Ile		
225				230					235					240			
Thr	Gly	Asn	Thr	Ser	Leu	Val	Phe	Ser	Glu	Asn	Ser	Val	Thr	Ala	Thr		
				245					250					255			
Ala	Gly	Asn	Gly	Gly	Ala	Leu	Ser	Gly	Asp	Ala	Asp	Val	Thr	Ile	Ser		
		260						265					270				
Gly	Asn	Gln	Ser	Val	Thr	Phe	Ser	Gly	Asn	Gln	Ala	Val	Ala	Asn	Gly		
	275					280					285						
Gly	Ala	Ile	Tyr	Ala	Lys	Lys	Leu	Thr	Leu	Ala	Ser	Gly	Gly	Gly	Gly		
	290				295						300						
Gly	Ile	Ser	Phe	Ser	Asn	Asn	Ile	Val	Gln	Gly	Thr	Thr	Ala	Gly	Asn		
305					310				315					320			
Gly	Gly	Ala	Ile	Ser	Ile	Leu	Ala	Ala	Gly	Glu	Cys	Ser	Leu	Ser	Ala		
			325					330					335				
Glu	Ala	Gly	Asp	Ile	Thr	Phe	Asn	Gly	Asn	Ala	Ile	Val	Ala	Thr	Thr		
			340					345					350				
Pro	Gln	Thr	Thr	Lys	Arg	Asn	Ser	Ile	Asp	Ile	Gly	Ser	Thr	Ala	Lys		
		355				360						365					
Ile	Thr	Asn	Leu	Arg	Ala	Ile	Ser	Gly	His	Ser	Ile	Phe	Phe	Tyr	Asp		
	370					375					380						

Pro Ile Thr Ala Asn Thr Ala Ala Asp Ser Thr Asp Thr Leu Asn Leu  
 385 390 395 400  
 Asn Lys Ala Asp Ala Gly Asn Ser Thr Asp Tyr Ser Gly Ser Ile Val  
 405 410 415  
 Phe Ser Gly Glu Lys Leu Ser Glu Asp Glu Ala Lys Val Ala Asp Asn  
 420 425 430  
 Leu Thr Ser Thr Leu Lys Gln Pro Val Thr Leu Thr Ala Gly Asn Leu  
 435 440 445  
 Val Leu Lys Arg Gly Val Thr Leu Asp Thr Lys Gly Phe Thr Gln Thr  
 450 455 460  
 Ala Gly Ser Ser Val Ile Met Asp Ala Gly Thr Thr Leu Lys Ala Ser  
 465 470 475 480  
 Thr Glu Glu Val Thr Leu Thr Gly Leu Ser Ile Pro Val Asp Ser Leu  
 485 490 495  
 Gly Glu Gly Lys Lys Val Val Ile Ala Ala Ser Ala Ala Ser Lys Asn  
 500 505 510  
 Val Ala Leu Ser Gly Pro Ile Leu Leu Leu Asp Asn Gln Gly Asn Ala  
 515 520 525  
 Tyr Glu Asn His Asp Leu Gly Lys Thr Gln Asp Phe Ser Phe Val Gln  
 530 535 540  
 Leu Ser Ala Leu Gly Thr Ala Thr Thr Thr Asp Val Pro Ala Val Pro  
 545 550 555 560  
 Thr Val Ala Thr Pro Thr His Tyr Gly Tyr Gln Gly Thr Trp Gly Met  
 565 570 575  
 Thr Trp Val Asp Asp Thr Ala Ser Thr Pro Lys Thr Lys Thr Ala Thr  
 580 585 590  
 Leu Ala Trp Thr Asn Thr Gly Tyr Leu Pro Asn Pro Glu Arg Gln Gly  
 595 600 605  
 Pro Leu Val Pro Asn Ser Leu Trp Gly Ser Phe Ser Asp Ile Gln Ala  
 610 615 620  
 Ile Gln Gly Val Ile Glu Arg Ser Ala Leu Thr Leu Cys Ser Asp Arg  
 625 630 635 640  
 Gly Phe Trp Ala Ala Gly Val Ala Asn Phe Leu Asp Lys Asp Lys Lys  
 645 650 655  
 Gly Glu Lys Arg Lys Tyr Arg His Lys Ser Gly Gly Tyr Ala Ile Gly  
 660 665 670  
 Gly Ala Ala Gln Thr Cys Ser Glu Asn Leu Ile Ser Phe Ala Phe Cys  
 675 680 685  
 Gln Leu Phe Gly Ser Asp Lys Asp Phe Leu Val Ala Lys Asn His Thr  
 690 695 700

Asp Thr Tyr Ala Gly Ala Phe Tyr Ile Gln His Ile Thr Glu Cys Ser  
 705 710 715 720  
 Gly Phe Ile Gly Cys Leu Leu Asp Lys Leu Pro Gly Ser Trp Ser His  
 725 730 735  
 Lys Pro Leu Val Leu Glu Gly Gln Leu Ala Tyr Ser His Val Ser Asn  
 740 745 750  
 Asp Leu Lys Thr Lys Tyr Thr Ala Tyr Pro Glu Val Lys Gly Ser Trp  
 755 760 765  
 Gly Asn Asn Ala Phe Asn Met Met Leu Gly Ala Ser Ser His Ser Tyr  
 770 775 780  
 Pro Glu Tyr Leu His Cys Phe Asp Thr Tyr Ala Pro Tyr Ile Lys Leu  
 785 790 795 800  
 Asn Leu Thr Tyr Ile Arg Gln Asp Ser Phe Ser Glu Lys Gly Thr Glu  
 805 810 815  
 Gly Arg Ser Phe Asp Asp Ser Asn Leu Phe Asn Leu Ser Leu Pro Ile  
 820 825 830  
 Gly Val Lys Phe Glu Lys Phe Ser Asp Cys Asn Asp Phe Ser Tyr Asp  
 835 840 845  
 Leu Thr Leu Ser Tyr Val Pro Asp Leu Ile Arg Asn Asp Pro Lys Cys  
 850 855 860  
 Thr Thr Ala Leu Val Ile Ser Gly Ala Ser Trp Glu Thr Tyr Ala Asn  
 865 870 875 880  
 Asn Leu Ala Arg Gln Ala Leu Gln Val Arg Ala Gly Ser His Tyr Ala  
 885 890 895  
 Phe Ser Pro Met Phe Glu Val Leu Gly Gln Phe Val Phe Glu Val Arg  
 900 905 910  
 Gly Ser

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1200
- (D) OTHER INFORMATION:



## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GAT Asp 1	CCT Pro	AAA Lys	AAT Asn	AAA Lys 5	GAG Glu	TAC Tyr	ACA Thr	GGG Gly	ACC Thr 10	ATA Ile	CTC Leu	TTT Phe	TCT Ser	GGA Gly 15	GAA Glu	48
AAG Lys	AGT Ser	CTA Leu	GCA Ala 20	AAC Asn	GAT Asp	CCT Pro	AGG Arg	GAT Asp 25	TTT Phe	AAA Lys	TCT Ser	ACA Thr	ATC Ile 30	CCT Pro	CAG Gln	96
AAC Asn	GTC Val	AAC Asn 35	CTG Leu	TCT Ser	GCA Ala	GGA Gly	TAC Tyr 40	TTA Leu	GTT Val	ATT Ile	AAA Lys	GAG Glu 45	GGG Gly	GCC Ala	GAA Glu	144
GTC Val	ACA Thr 50	GTT Val	TCA Ser	AAA Lys	TTC Phe	ACG Thr 55	CAG Gln	TCT Ser	CCA Pro	GGA Gly	TCG Ser 60	CAT His	TTA Leu	GTT Val	TTA Leu	192
GAT Asp 65	TTA Leu	GGA Gly	ACC Thr	AAA Lys	CTG Leu 70	ATA Ile	GCC Ala	TCT Ser	AAG Lys	GAA Glu 75	GAC Asp	ATT Ile	GCC Ala	ATC Ile	ACA Thr 80	240
GGC Gly	CTC Leu	GCG Ala	ATA Ile	GAT Asp 85	ATA Ile	GAT Asp	AGC Ser	TTA Leu	AGC Ser 90	TCA Ser	TCC Ser	TCA Ser	ACA Thr	GCA Ala 95	GCT Ala	288
GTT Val	ATT Ile	AAA Lys	GCA Ala 100	AAC Asn	ACC Thr	GCA Ala	AAT Asn	AAA Lys 105	CAG Gln	ATA Ile	TCC Ser	GTG Val	ACG Thr 110	GAC Asp	TCT Ser	336
ATA Ile	GAA Glu 115	CTT Leu	ATC Ile	TCG Ser	CCT Pro	ACT Thr	GGC Gly 120	AAT Asn	GCC Ala	TAT Tyr	GAA Glu	GAT Asp 125	CTC Leu	AGA Arg	ATG Met	384
AGA Arg 130	AAT Asn	TCA Ser	CAG Gln	ACG Thr	TTC Phe	CCT Pro	CTG Leu 135	CTC Leu	TCT Ser	TTA Leu	GAG Glu 140	CCT Pro	GGA Gly	GCC Ala	GGG Gly	432
GGT Gly 145	AGT Ser	GTG Val	ACT Thr	GTA Val	ACT Thr 150	GCT Ala	GGA Gly	GAT Asp	TTC Phe	CTA Leu 155	CCG Pro	GTA Val	AGT Ser	CCC Pro	CAT His 160	480
TAT Tyr	GGT Gly	TTT Phe	CAA Gln 165	GGC Gly	AAT Asn	TGG Trp	AAA Lys	TTA Leu 170	GCT Ala	TGG Trp	ACA Thr	GGA Gly	ACT Thr	GGA Gly 175	AAC Asn	528
AAA Lys	GTT Val	GGA Gly	GAA Glu 180	TTC Phe	TTC Phe	TGG Trp	GAT Asp	AAA Lys 185	ATA Ile	AAT Asn	TAT Tyr	AAG Lys	CCT Pro	AGA Arg	CCT Pro	576
GAA Glu	AAA Lys	GAA Glu 195	GGA Gly	AAT Asn	TTA Leu	GTT Val	CCT Pro	AAT Asn 200	ATC Ile	TTG Leu	TGG Trp	GGG Gly 205	AAT Asn	GCT Ala	GTA Val	624
AAT Asn 210	GTC Val	AGA Arg	TCC Ser	TTA Leu	ATG Met	CAG Gln 215	GTT Val	CAA Gln	GAG Glu	ACC Thr	CAT His 220	GCA Ala	TCG Ser	AGC Ser	TTA Leu	672

CAG	ACA	GAT	CGA	GGG	CTG	TGG	ATC	GAT	GGA	ATT	GGG	AAT	TTC	TTC	CAT	720
Gln	Thr	Asp	Arg	Gly	Leu	Trp	Ile	Asp	Gly	Ile	Gly	Asn	Phe	Phe	His	
225					230					235					240	
GTA	TCT	GCC	TCC	GAA	GAC	AAT	ATA	AGG	TAC	CGT	CAT	AAC	AGC	GGT	GGA	768
Val	Ser	Ala	Ser	Glu	Asp	Asn	Ile	Arg	Tyr	Arg	His	Asn	Ser	Gly	Gly	
				245					250					255		
TAT	GTT	CTA	TCT	GTA	AAT	AAT	GAG	ATC	ACA	CCT	AAG	CAC	TAT	ACT	TCG	816
Tyr	Val	Leu	Ser	Val	Asn	Asn	Glu	Ile	Thr	Pro	Lys	His	Tyr	Thr	Ser	
			260					265					270			
ATG	GCA	TTT	TCC	CAA	CTC	TTT	AGT	AGA	GAC	AAA	GAC	TAT	GCG	GTT	TCC	864
Met	Ala	Phe	Ser	Gln	Leu	Phe	Ser	Arg	Asp	Lys	Asp	Tyr	Ala	Val	Ser	
		275						280				285				
AAC	AAC	GAA	TAC	AGA	ATG	TAT	TTA	GGA	TCG	TAT	CTC	TAT	CAA	TAT	ACA	912
Asn	Asn	Glu	Tyr	Arg	Met	Tyr	Leu	Gly	Ser	Tyr	Leu	Tyr	Gln	Tyr	Thr	
		290					295					300				
ACC	TCC	CTA	GGG	AAT	ATT	TTC	CGT	TAT	GCT	TCG	CGT	AAC	CCT	AAT	GTA	960
Thr	Ser	Leu	Gly	Asn	Ile	Phe	Arg	Tyr	Ala	Ser	Arg	Asn	Pro	Asn	Val	
305					310					315					320	
AAC	GTC	GGG	ATT	CTC	TCA	AGA	AGG	TTT	CTT	CAA	AAT	CCT	CTT	ATG	ATT	1008
Asn	Val	Gly	Ile	Leu	Ser	Arg	Arg	Phe	Leu	Gln	Asn	Pro	Leu	Met	Ile	
				325					330					335		
TTT	CAT	TTT	TTG	TGT	GCT	TAT	GGT	CAT	GCC	ACC	AAT	GAT	ATG	AAA	ACA	1056
Phe	His	Phe	Leu	Cys	Ala	Tyr	Gly	His	Ala	Thr	Asn	Asp	Met	Lys	Thr	
			340					345					350			
GAC	TAC	GCA	AAT	TTC	CCT	ATG	GTG	AAA	AAC	AGC	TGG	AGA	AAC	AAT	TGT	1104
Asp	Tyr	Ala	Asn	Phe	Pro	Met	Val	Lys	Asn	Ser	Trp	Arg	Asn	Asn	Cys	
		355					360					365				
TGG	GCT	ATA	AAA	TGC	GGA	GGG	AGC	ATG	CCT	CTA	TTG	GTA	TTT	GAA	AAC	1152
Trp	Ala	Ile	Lys	Cys	Gly	Gly	Ser	Met	Pro	Leu	Leu	Val	Phe	Glu	Asn	
		370				375						380				
GGA	AAA	CTT	TTC	CAA	GGT	GCC	ATC	CCA	TTT	ATG	AAA	CTA	CAA	TTA	GTT	1200
Gly	Lys	Leu	Phe	Gln	Gly	Ala	Ile	Pro	Phe	Met	Lys	Leu	Gln	Leu	Val	
385					390					395					400	

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 400 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (v) FRAGMENT TYPE: internal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Asp Pro Lys Asn Lys Glu Tyr Thr Gly Thr Ile Leu Phe Ser Gly Glu  
 1 5 10 15  
 Lys Ser Leu Ala Asn Asp Pro Arg Asp Phe Lys Ser Thr Ile Pro Gln  
 20 25 30  
 Asn Val Asn Leu Ser Ala Gly Tyr Leu Val Ile Lys Glu Gly Ala Glu  
 35 40 45  
 Val Thr Val Ser Lys Phe Thr Gln Ser Pro Gly Ser His Leu Val Leu  
 50 55 60  
 Asp Leu Gly Thr Lys Leu Ile Ala Ser Lys Glu Asp Ile Ala Ile Thr  
 65 70 75 80  
 Gly Leu Ala Ile Asp Ile Asp Ser Leu Ser Ser Ser Thr Ala Ala  
 85 90 95  
 Val Ile Lys Ala Asn Thr Ala Asn Lys Gln Ile Ser Val Thr Asp Ser  
 100 105 110  
 Ile Glu Leu Ile Ser Pro Thr Gly Asn Ala Tyr Glu Asp Leu Arg Met  
 115 120 125  
 Arg Asn Ser Gln Thr Phe Pro Leu Leu Ser Leu Glu Pro Gly Ala Gly  
 130 135 140  
 Gly Ser Val Thr Val Thr Ala Gly Asp Phe Leu Pro Val Ser Pro His  
 145 150 155 160  
 Tyr Gly Phe Gln Gly Asn Trp Lys Leu Ala Trp Thr Gly Thr Gly Asn  
 165 170 175  
 Lys Val Gly Glu Phe Phe Trp Asp Lys Ile Asn Tyr Lys Pro Arg Pro  
 180 185 190  
 Glu Lys Glu Gly Asn Leu Val Pro Asn Ile Leu Trp Gly Asn Ala Val  
 195 200 205  
 Asn Val Arg Ser Leu Met Gln Val Gln Glu Thr His Ala Ser Ser Leu  
 210 215 220  
 Gln Thr Asp Arg Gly Leu Trp Ile Asp Gly Ile Gly Asn Phe Phe His  
 225 230 235 240  
 Val Ser Ala Ser Glu Asp Asn Ile Arg Tyr Arg His Asn Ser Gly Gly  
 245 250 255  
 Tyr Val Leu Ser Val Asn Asn Glu Ile Thr Pro Lys His Tyr Thr Ser  
 260 265 270  
 Met Ala Phe Ser Gln Leu Phe Ser Arg Asp Lys Asp Tyr Ala Val Ser  
 275 280 285  
 Asn Asn Glu Tyr Arg Met Tyr Leu Gly Ser Tyr Leu Tyr Gln Tyr Thr  
 290 295 300

```

Thr Ser Leu Gly Asn Ile Phe Arg Tyr Ala Ser Arg Asn Pro Asn Val
305                      310                      315                      320

Asn Val Gly Ile Leu Ser Arg Arg Phe Leu Gln Asn Pro Leu Met Ile
                      325                      330                      335

Phe His Phe Leu Cys Ala Tyr Gly His Ala Thr Asn Asp Met Lys Thr
                      340                      345                      350

Asp Tyr Ala Asn Phe Pro Met Val Lys Asn Ser Trp Arg Asn Asn Cys
                      355                      360                      365

Trp Ala Ile Lys Cys Gly Gly Ser Met Pro Leu Leu Val Phe Glu Asn
370                      375                      380

Gly Lys Leu Phe Gln Gly Ala Ile Pro Phe Met Lys Leu Gln Leu Val
385                      390                      395                      400

```

## (2) INFORMATION FOR SEQ ID NO:29:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1830 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 1...1830
- (D) OTHER INFORMATION:

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

```

GAT CTC ACA TTA GGG AGT CGT GAC AGT TAT AAT GGT GAT ACA AGC ACC      48
Asp Leu Thr Leu Gly Ser Arg Asp Ser Tyr Asn Gly Asp Thr Ser Thr
  1                      5                      10                      15

ACA GAA TTT ACT CCT AAA GCG GCA ACT TCT GAT GCT AGT GGC ACG ACC      96
Thr Glu Phe Thr Pro Lys Ala Ala Thr Ser Asp Ala Ser Gly Thr Thr
                      20                      25                      30

TAT ATT CTC GAT GGG GAT GTC TCG ATA AGC CAA GCA GGG AAA CAA ACG     144
Tyr Ile Leu Asp Gly Asp Val Ser Ile Ser Gln Ala Gly Lys Gln Thr
                      35                      40                      45

AGC TTA ACC ACA AGT TGT TTT TCT AAC ACT GCA GGA AAT CTT ACC TTC     192
Ser Leu Thr Thr Ser Cys Phe Ser Asn Thr Ala Gly Asn Leu Thr Phe
                      50                      55                      60

TTA GGG AAC GGA TTT TCT CTT CAT TTT GAC AAT ATT ATT TCG TCT ACT     240
Leu Gly Asn Gly Phe Ser Leu His Phe Asp Asn Ile Ile Ser Ser Thr
                      65                      70                      75                      80

GTT GCA GGT GTT GTT GTT AGC AAT ACA GCA GCT TCT GGG ATT ACG AAA     288
Val Ala Gly Val Val Val Ser Asn Thr Ala Ala Ser Gly Ile Thr Lys

```

95

TTC Phe	TCA Ser	GGA Gly	TTT Phe 100	TCA Ser	ACT Thr	CTT Leu	CGG Arg 105	ATG Met	CTT Leu	GCA Ala	GCT Ala	CCT Pro 110	AGG Arg	ACC Thr	ACA Thr	336
GGT Gly	AAA Lys	GGA Gly 115	GCC Ala	ATT Ile	AAA Lys	ATT Ile	ACC Thr 120	GAT Asp	GGT Gly	CTG Leu	GTG Val	TTT Phe 125	GAG Glu	AGT Ser	ATA Ile	384
GGG Gly	AAT Asn 130	CTT Leu	GAT Asp	CCG Pro	ATT Ile	ACT Thr 135	GTA Val	ACA Thr	GGA Gly	TCG Ser 140	ACA Thr	TCT Ser	GTT Val	GCT Ala	GAT Asp	432
GCT Ala 145	CTC Leu	AAT Asn	ATT Ile	AAT Asn	AGC Ser 150	CCT Pro	GAT Asp	ACT Thr	GGA Gly	GAT Asp 155	AAC Asn	AAA Lys	GAG Glu	TAT Tyr	ACG Thr 160	480
GGA Gly	ACC Thr	ATA Ile	GTC Val 165	TTT Phe 165	TCT Ser	GGA Gly	GAG Glu	AAG Lys	CTC Leu 170	ACG Thr	GAG Glu	GCA Ala	GAA Glu	GCT Ala 175	AAA Lys	528
GAT Asp	GAG Glu	AAG Lys	AAC Asn 180	CGC Arg	ACT Thr	TCT Ser	AAA Lys 185	TTA Leu	CTT Leu	CAA Gln	AAT Asn	GTT Val 190	GCT Ala	TTT Phe	AAA Lys	576
AAT Asn	GGG Gly	ACT Thr 195	GTA Val	GTT Val	TTA Leu	AAA Lys	GGT Gly 200	GAT Asp	GTC Val	GTT Val	TTA Leu	AGT Ser 205	GCG Ala	AAC Asn	GGT Gly	624
TTC Phe 210	TCT Ser	CAG Gln	GAT Asp	GCA Ala	AAC Asn	TCT Ser 215	AAG Lys	TTG Leu	ATT Ile	ATG Met	GAT Asp 220	TTA Leu	GGG Gly	ACG Thr	TCG Ser	672
TTG Leu 225	GTT Val	GCA Ala	AAC Asn	ACC Thr	GAA Glu 230	AGT Ser	ATC Ile	GAG Glu	TTA Leu	ACG Thr 235	AAT Asn	TTG Leu	GAA Glu	ATT Ile	AAT Asn 240	720
ATA Ile	GAC Asp	TCT Ser	CTC Leu	AGG Arg 245	AAC Asn	GGG Gly	AAA Lys	AAG Lys	ATA Ile 250	AAA Lys	CTC Leu	AGT Ser	GCT Ala	GCC Ala 255	ACA Thr	768
GCT Ala	CAG Gln	AAA Lys	GAT Asp 260	ATT Ile	CGT Arg	ATA Ile	GAT Asp 265	CGT Arg	CCT Pro	GTT Val	GTA Val	CTG Leu	GCA Ala 270	ATT Ile	AGC Ser	816
GAT Asp	GAG Glu	AGT Ser 275	TTT Phe	TAT Tyr	CAA Gln	AAT Asn	GGC Gly 280	TTT Phe	TTG Leu	AAT Asn	GAG Glu	GAC Asp 285	CAT His	TCC Ser	TAT Tyr	864
GAT Asp	GGG Gly 290	ATT Ile	CTT Leu	GAG Glu	TTA Leu	GAT Asp 295	GCT Ala	GGG Gly	AAA Lys	GAC Asp	ATC Ile 300	GTG Val	ATT Ile	TCT Ser	GCA Ala	912
GAT Asp 305	TCT Ser	CGC Arg	AGT Ser	ATA Ile	GAT Asp 310	GCT Ala	GTA Val	CAA Gln	TCT Ser	CCG Pro 315	TAT Tyr	GGC Gly	TAT Tyr	CAG Gln	GGA Gly 320	960

AAG	TGG	ACG	ATC	AAT	TGG	TCT	ACT	GAT	GAT	AAG	AAA	GCT	ACG	GTT	TCT	1008
Lys	Trp	Thr	Ile	Asn	Trp	Ser	Thr	Asp	Asp	Lys	Lys	Ala	Thr	Val	Ser	
				325					330					335		
TGG	GCG	AAG	CAG	AGT	TTT	AAT	CCC	ACT	GCT	GAG	CAG	GAG	GCT	CCG	TTA	1056
Trp	Ala	Lys	Gln	Ser	Phe	Asn	Pro	Thr	Ala	Glu	Gln	Glu	Ala	Pro	Leu	
			340					345					350			
GTT	CCT	AAT	CTT	CTT	TGG	GGT	TCT	TTT	ATA	GAT	GTT	CGT	TCC	TTC	CAG	1104
Val	Pro	Asn	Leu	Leu	Trp	Gly	Ser	Phe	Ile	Asp	Val	Arg	Ser	Phe	Gln	
		355					360					365				
AAT	TTT	ATA	GAG	CTA	GGT	ACT	GAA	GGT	GCT	CCT	TAC	GAA	AAG	AGA	TTT	1152
Asn	Phe	Ile	Glu	Leu	Gly	Thr	Glu	Gly	Ala	Pro	Tyr	Glu	Lys	Arg	Phe	
	370					375					380					
TGG	GTT	GCA	GGC	ATT	TCC	AAT	GTT	TTG	CAT	AGG	AGC	GGT	CGT	GAA	AAT	1200
Trp	Val	Ala	Gly	Ile	Ser	Asn	Val	Leu	His	Arg	Ser	Gly	Arg	Glu	Asn	
385					390					395					400	
CAA	AGG	AAA	TTC	CGT	CAT	GTG	AGT	GGA	GGT	GCT	GTA	GTA	GGT	GCT	AGC	1248
Gln	Arg	Lys	Phe	Arg	His	Val	Ser	Gly	Gly	Ala	Val	Val	Gly	Ala	Ser	
				405					410					415		
ACG	AGG	ATG	CCG	GGT	GGT	GAT	ACC	TTG	TCT	CTG	GGT	TTT	GCT	CAG	CTC	1296
Thr	Arg	Met	Pro	Gly	Gly	Asp	Thr	Leu	Ser	Leu	Gly	Phe	Ala	Gln	Leu	
			420					425					430			
TTT	GCG	CGT	GAC	AAA	GAC	TAC	TTT	ATG	AAT	ACC	AAT	TTC	GCA	AAG	ACC	1344
Phe	Ala	Arg	Asp	Lys	Asp	Tyr	Phe	Met	Asn	Thr	Asn	Phe	Ala	Lys	Thr	
		435					440					445				
TAC	GCA	GGA	TCT	TTA	CGT	TTG	CAG	CAC	GAT	GCT	TCC	CTA	TAC	TCT	GTG	1392
Tyr	Ala	Gly	Ser	Leu	Arg	Leu	Gln	His	Asp	Ala	Ser	Leu	Tyr	Ser	Val	
	450					455					460					
GTG	AGT	ATC	CTT	TTA	GGA	GAG	GGA	GGA	CTC	CGC	GAG	ATC	CTG	TTG	CCT	1440
Val	Ser	Ile	Leu	Leu	Gly	Glu	Gly	Gly	Leu	Arg	Glu	Ile	Leu	Leu	Pro	
465					470				475						480	
TAT	GTT	TCC	AAT	ACT	CTG	CCG	TGC	TCT	TTC	TAT	GGG	CAG	CTT	AGC	TAC	1488
Tyr	Val	Ser	Asn	Thr	Leu	Pro	Cys	Ser	Phe	Tyr	Gly	Gln	Leu	Ser	Tyr	
			485					490						495		
GGC	CAT	ACG	GAT	CAT	CGC	ATG	AAG	ACC	GAG	TCT	CTA	CCC	CCC	CCC	CCC	1536
Gly	His	Thr	Asp	His	Arg	Met	Lys	Thr	Glu	Ser	Leu	Pro	Pro	Pro	Pro	
			500					505				510				
CCG	ACG	CTC	TCG	ACG	GAT	CAT	ACT	TCT	TGG	GGA	GGA	TAT	GTC	TGG	GCT	1584
Pro	Thr	Leu	Ser	Thr	Asp	His	Thr	Ser	Trp	Gly	Gly	Tyr	Val	Trp	Ala	
		515					520					525				
GGA	GAG	CTG	GGA	ACT	CGA	GTT	GCT	GTT	GAA	AAT	ACC	AGC	GGC	AGA	GGA	1632
Gly	Glu	Leu	Gly	Thr	Arg	Val	Ala	Val	Glu	Asn	Thr	Ser	Gly	Arg	Gly	
	530					535					540					
TTT	TTC	CGA	GAG	TAC	ACT	CCA	TTT	GTA	AAA	GTC	CAA	GCT	GTT	TAC	TCG	1680
Phe	Phe	Arg	Glu	Tyr	Thr	Pro	Phe	Val	Lys	Val	Gln	Ala	Val	Tyr	Ser	

545	550	555	560	
CGC CAA GAT AGC TTT GTT GAA CTA GGA GCT ATC AGT CGT GAT TTT AGT				1728
Arg Gln Asp Ser Phe Val Glu Leu Gly Ala Ile Ser Arg Asp Phe Ser				
565	570		575	
GAT TCG CAT CTT TAT AAC CTT GCG ATT CCT CTT GGA ATC AAG TTA GAG				1776
Asp Ser His Leu Tyr Asn Leu Ala Ile Pro Leu Gly Ile Lys Leu Glu				
580	585		590	
AAA CGG TTT GCA GAG CAA TAT TAT CAT GTT GTT GCG ATG TAT TCT CCA				1824
Lys Arg Phe Ala Glu Gln Tyr Tyr His Val Val Ala Met Tyr Ser Pro				
595	600		605	
GAT GTT				1830
Asp Val				
610				

## (2) INFORMATION FOR SEQ ID NO:30:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 610 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (v) FRAGMENT TYPE: internal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Asp	Leu	Thr	Leu	Gly	Ser	Arg	Asp	Ser	Tyr	Asn	Gly	Asp	Thr	Ser	Thr
1				5					10					15	
Thr	Glu	Phe	Thr	Pro	Lys	Ala	Ala	Thr	Ser	Asp	Ala	Ser	Gly	Thr	Thr
			20					25					30		
Tyr	Ile	Leu	Asp	Gly	Asp	Val	Ser	Ile	Ser	Gln	Ala	Gly	Lys	Gln	Thr
		35				40						45			
Ser	Leu	Thr	Thr	Ser	Cys	Phe	Ser	Asn	Thr	Ala	Gly	Asn	Leu	Thr	Phe
		50				55					60				
Leu	Gly	Asn	Gly	Phe	Ser	Leu	His	Phe	Asp	Asn	Ile	Ile	Ser	Ser	Thr
	65				70					75					80
Val	Ala	Gly	Val	Val	Val	Ser	Asn	Thr	Ala	Ala	Ser	Gly	Ile	Thr	Lys
				85					90					95	
Phe	Ser	Gly	Phe	Ser	Thr	Leu	Arg	Met	Leu	Ala	Ala	Pro	Arg	Thr	Thr
			100					105					110		
Gly	Lys	Gly	Ala	Ile	Lys	Ile	Thr	Asp	Gly	Leu	Val	Phe	Glu	Ser	Ile
		115					120					125			
Gly	Asn	Leu	Asp	Pro	Ile	Thr	Val	Thr	Gly	Ser	Thr	Ser	Val	Ala	Asp

130					135					140					
Ala	Leu	Asn	Ile	Asn	Ser	Pro	Asp	Thr	Gly	Asp	Asn	Lys	Glu	Tyr	Thr
145					150					155					160
Gly	Thr	Ile	Val	Phe	Ser	Gly	Glu	Lys	Leu	Thr	Glu	Ala	Glu	Ala	Lys
				165					170					175	
Asp	Glu	Lys	Asn	Arg	Thr	Ser	Lys	Leu	Leu	Gln	Asn	Val	Ala	Phe	Lys
			180					185					190		
Asn	Gly	Thr	Val	Val	Leu	Lys	Gly	Asp	Val	Val	Leu	Ser	Ala	Asn	Gly
		195					200					205			
Phe	Ser	Gln	Asp	Ala	Asn	Ser	Lys	Leu	Ile	Met	Asp	Leu	Gly	Thr	Ser
	210					215					220				
Leu	Val	Ala	Asn	Thr	Glu	Ser	Ile	Glu	Leu	Thr	Asn	Leu	Glu	Ile	Asn
225				230						235					240
Ile	Asp	Ser	Leu	Arg	Asn	Gly	Lys	Lys	Ile	Lys	Leu	Ser	Ala	Ala	Thr
				245					250					255	
Ala	Gln	Lys	Asp	Ile	Arg	Ile	Asp	Arg	Pro	Val	Val	Leu	Ala	Ile	Ser
			260					265					270		
Asp	Glu	Ser	Phe	Tyr	Gln	Asn	Gly	Phe	Leu	Asn	Glu	Asp	His	Ser	Tyr
		275					280					285			
Asp	Gly	Ile	Leu	Glu	Leu	Asp	Ala	Gly	Lys	Asp	Ile	Val	Ile	Ser	Ala
	290					295					300				
Asp	Ser	Arg	Ser	Ile	Asp	Ala	Val	Gln	Ser	Pro	Tyr	Gly	Tyr	Gln	Gly
305				310						315					320
Lys	Trp	Thr	Ile	Asn	Trp	Ser	Thr	Asp	Asp	Lys	Lys	Ala	Thr	Val	Ser
				325					330					335	
Trp	Ala	Lys	Gln	Ser	Phe	Asn	Pro	Thr	Ala	Glu	Gln	Glu	Ala	Pro	Leu
			340					345					350		
Val	Pro	Asn	Leu	Leu	Trp	Gly	Ser	Phe	Ile	Asp	Val	Arg	Ser	Phe	Gln
		355					360					365			
Asn	Phe	Ile	Glu	Leu	Gly	Thr	Glu	Gly	Ala	Pro	Tyr	Glu	Lys	Arg	Phe
	370					375					380				
Trp	Val	Ala	Gly	Ile	Ser	Asn	Val	Leu	His	Arg	Ser	Gly	Arg	Glu	Asn
385				390						395					400
Gln	Arg	Lys	Phe	Arg	His	Val	Ser	Gly	Gly	Ala	Val	Val	Gly	Ala	Ser
			405						410					415	
Thr	Arg	Met	Pro	Gly	Gly	Asp	Thr	Leu	Ser	Leu	Gly	Phe	Ala	Gln	Leu
			420					425					430		
Phe	Ala	Arg	Asp	Lys	Asp	Tyr	Phe	Met	Asn	Thr	Asn	Phe	Ala	Lys	Thr
		435					440					445			



Tyr Ala Gly Ser Leu Arg Leu Gln His Asp Ala Ser Leu Tyr Ser Val  
 450 455 460  
 Val Ser Ile Leu Leu Gly Glu Gly Gly Leu Arg Glu Ile Leu Leu Pro  
 465 470 475 480  
 Tyr Val Ser Asn Thr Leu Pro Cys Ser Phe Tyr Gly Gln Leu Ser Tyr  
 485 490 495  
 Gly His Thr Asp His Arg Met Lys Thr Glu Ser Leu Pro Pro Pro Pro  
 500 505 510  
 Pro Thr Leu Ser Thr Asp His Thr Ser Trp Gly Gly Tyr Val Trp Ala  
 515 520 525  
 Gly Glu Leu Gly Thr Arg Val Ala Val Glu Asn Thr Ser Gly Arg Gly  
 530 535 540  
 Phe Phe Arg Glu Tyr Thr Pro Phe Val Lys Val Gln Ala Val Tyr Ser  
 545 550 555 560  
 Arg Gln Asp Ser Phe Val Glu Leu Gly Ala Ile Ser Arg Asp Phe Ser  
 565 570 575  
 Asp Ser His Leu Tyr Asn Leu Ala Ile Pro Leu Gly Ile Lys Leu Glu  
 580 585 590  
 Lys Arg Phe Ala Glu Gln Tyr Tyr His Val Val Ala Met Tyr Ser Pro  
 595 600 605  
 Asp Val  
 610